

AD-A233 275

MENTATION PAGE

Form Approved
OMB No. 0704-0188

is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and reviewing the collection of information, sending comments regarding this burden estimate or any other aspect of this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Avenue, Suite 1204, Washington, DC 20540-6001, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 31 January '91		3. REPORT TYPE AND DATES COVERED Annual Technical Report 1 Feb.90-31 Jan91	
4. TITLE AND SUBTITLE CURVATURE ESTIMATION IN ORIENTATION SELECTION w. / Atch				5. FUNDING NUMBERS AFOSR-89-0260 2313/A8 121102F	
6. AUTHOR(S) Professor Steven W. Zucker Professor Max S. Cynader				8. PERFORMING ORGANIZATION REPORT NUMBER CIM 90-19	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) McGill University 3550 University Street Montreal, Quebec, Canada H3A 2A7				10. SPONSORING / MONITORING AGENCY REPORT NUMBER AFOSR-R- 1 0118	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR Building 410 Bolling Air Force Base, D.C. 20332-6448 U.S.A.					
11. SUPPLEMENTARY NOTES					
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				12b. DISTRIBUTION CODE DTIC ELECTE MAR 12 1991 S E D	
13. ABSTRACT (Maximum 200 words) Abstract This research effort is concentrated on the computational neuroscience of early vision. Progress was made on the following problems: (i) development and simulation of our model for logical/linear interactions in visual cortical cells, application of it to several problems in visual psychophysics, and initial testing of physiological predictions from it; (ii) extension of our model of endstopped visual cortical neurons into motion and visual area MT; (iii) formal development of our model of shape, and initial psychophysical experiments inspired from it; and (iv) formal development of a model for inferring global curves from local (tangent) representations.					
14. SUBJECT TERMS				15. NUMBER OF PAGES	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified		18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified		19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	
20. LIMITATION OF ABSTRACT					

**Best
Available
Copy**

**Annual Technical Report
AFOSR Grant 89-0260
1 Feb 90 - 31 Jan 91**

Steven W. Zucker * Max S. Cynader * †

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Computer Vision and Robotics Laboratory
McGill Research Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada



* Fellow, Canadian Institute for Advanced Research.

*

† Dept. of Ophthalmology University of British Columbia Vancouver, B. C.

Postal Address: 3480 University Street, Montréal, Québec, Canada H3A 2A7

Telephone: (514) 398 6319 Telex: 05 268510 FAX: (514) 398 7348

Network Address: mrcim@larry.mrcim.mcgill.edu

91 3 06 157

Abstract

This research effort is concentrated on the computational neuroscience of early vision. Progress was made on the following problems: (i) development and simulation of our model for logical/linear interactions in visual cortical cells, application of it to several problems in visual psychophysics, and initial testing of physiological predictions from it; (ii) extension of our model of endstopped visual cortical neurons into motion and visual area MT; (iii) formal development of our model of shape, and initial psychophysical experiments inspired from it; and (iv) formal development of a model for inferring global curves from local (tangent) representations.

1. Introduction

Our second year of research support by AFOSR Grant 89-0260 has been an exciting one. Real progress has been made on several fronts, especially Computational Neuroscience, Theory, and Neurophysiology. Several papers have resulted from this research, and we discuss each project in turn. Since this grant is an interdisciplinary one, we begin with a project in which our computational modeling has already inspired physiological experiments.

2. Logical/Linear Modeling of Visual Cortical Cells

Our main area of collaboration continues to be the modeling of neurons in the visual cortex, but now with a concentration on which non-linearities might arise within them, what computational role these non-linearities could play in functional theories of vision, and how they can be studied physiologically. Most models of simple cells are linear (e.g., Gabor functions, differences of Gaussians, etc.), and these are combined (via rectification) into complex cells. Our work began with a computational motivation—that certain non-linearities are extremely important for separating nearby, but separate curves—and suggests a drastically different sort of non-linearity than has previously been investigated. In particular, we have developed a calculus for the logical combination of local evidence from subunits in which support is accumulated linearly, but incompatible evidence enters nonlinearly. This calculus is described in detail in the Technical Report in Appendix I. It leads to operators that appear linear for one class of stimuli but markedly nonlinear for others — we call these logical-linear operators. They exhibit dual advantages: they are considerably more stimulus-specific than purely linear operators, while more robust to incidental stimulus variation than logical operators. The viability of such operators as models of visual cortical neurons (e.g. simple cells) has been examined by comparing simulations of purely linear and logical-linear models to the responses of cortical neurons. Operators with properties specialised for spatial contours are examined with stimuli containing vernier offsets, interruptions, and opposite contrast segments. The results are consistent with the well-known “linear” properties (e.g., sensitivity to s.f. gratings) while exhibiting the nonlinear behaviour associated with high vernier sensitivities (Swindale and Cynader, 1989) and strong suppressive effects for opposite contrast segments (Hammond and MacKay, 1983, 1985). These latter results are described in Appendix II.

Our computational studies of logical/linear interactions have been so exciting that we have begun to develop physiological tests derived from them. These involve the development of novel stimuli in which short segments of opposite contrast bars are superimposed onto longer bars at preferred orientations, and then extended orthogonally. The end result is a cross composed of (say) a dark horizontal bar and a white vertical bar on a gray background. Such tests will permit us to understand significantly more about subunit interactions in receptive fields, and, if our predictions are substantiated, will force a rather different view of receptive fields in early vision.

The system for generating the relevant stimuli was completed this summer as a result of a student exchange between McGill University and the University of British Columbia, and initial testing has now begun under Cynader's direction. Data from the first two experiments is included in Figs. 1 and 2, and shows the response of the cell to a bar of preferred contrast with an opposite contrast blob at the center. The horizontal axis is contrast of the blob (same to the right, opposite to the left of center) vs. response. It is still too early to draw any conclusions from these data, but we expect to have a lot to report in the 6 month brief progress report.

3. Modeling of Endstopping and Curvature

Our other main area of collaboration has been the modeling of endstopped neurons in the visual cortex, with an examination of how well they might respond to curved stimuli. Recall that our model involves a rectified "difference" between the responses of two cells with oriented receptive fields, one large and the other small. Our model now includes complex cell components, and a very elaborate simulation of it has been completed. Two new papers, under review for *Visual Neuroscience*, are included in Appendices III and IV.

4. Motion and MT

We have begun to extend our theoretical efforts into the motion domain, and to extend our modeling efforts into area MT. After studying the treatment of discontinuities (Zucker, S.W., Iverson, L., and Hummel, R., Coherent compound motion: Corners and non-rigid configurations, *Neural Computation*, 1990, 2, 44 - 57), we have begun to address the question of how the visual system can obtain reliable estimates of the optical flow field and how this information can be used to obtain coarse estimates of shape. Substantial progress has been made; see Appendix V.

5. Toward a Computational Theory of Shape

We view the key bottleneck to high-level vision to be an analysis of shape, and thus have begun to address the question of how to decompose a contour into parts from a theoretical perspective. An early overview of the theory was included in last year's progress report; we now include recent theoretical results in Appendix VI.

In addition to our theoretical studies, we have also begun to examine psychophysical implications of the theory. In particular, we have discovered new classes of stimuli in which a single parameter controls a smooth variation from one shape into another; examples are included in Fig. 3 (can you locate the different "hourglass"?) where neck width is the parameter, and in Fig. 4,

(can you locate the "worm" in the field of sausages?) where phase angle is the parameter. Early data are also included, showing the time to locate the singular shape. Since the theory predicts which shapes are most similar, the psychophysical experiments are based on the tenet that more similar shapes should take longer to discriminate between. Clearly, although the shape parameters vary smoothly, the above prediction seems to be verified; there is an abrupt parameter value at which time increases substantially. Again, more psychophysics needs to be done, and results need to be quantified. This will be reported on at 6 months.

6. Summary and Conclusions

It should be clear from the above overview that our laboratories have been extremely busy, with projects in receptive field modeling, endstopping, motion, and shape all underway. There are in addition a number of other, more theoretical projects as well, such as the use of potential functions to fill the gulf between local and global visual representations (included as Appendix VII for completeness). While these projects may seem rather diverse when presented in this fashion, we invite you to visit our laboratories this year, to sample the activity more personally. This would give us the opportunity to discuss the larger view of vision that we have been evolving, and which inspires each of these more technical activities.

7. Supported Personnel

Four graduate students have been supported by this grant: Allan Dobbins (McGill), Benoit Dubuc (partial support; McGill), Lee Iverson (McGill), and Erica Strumpf (U.B.C.).

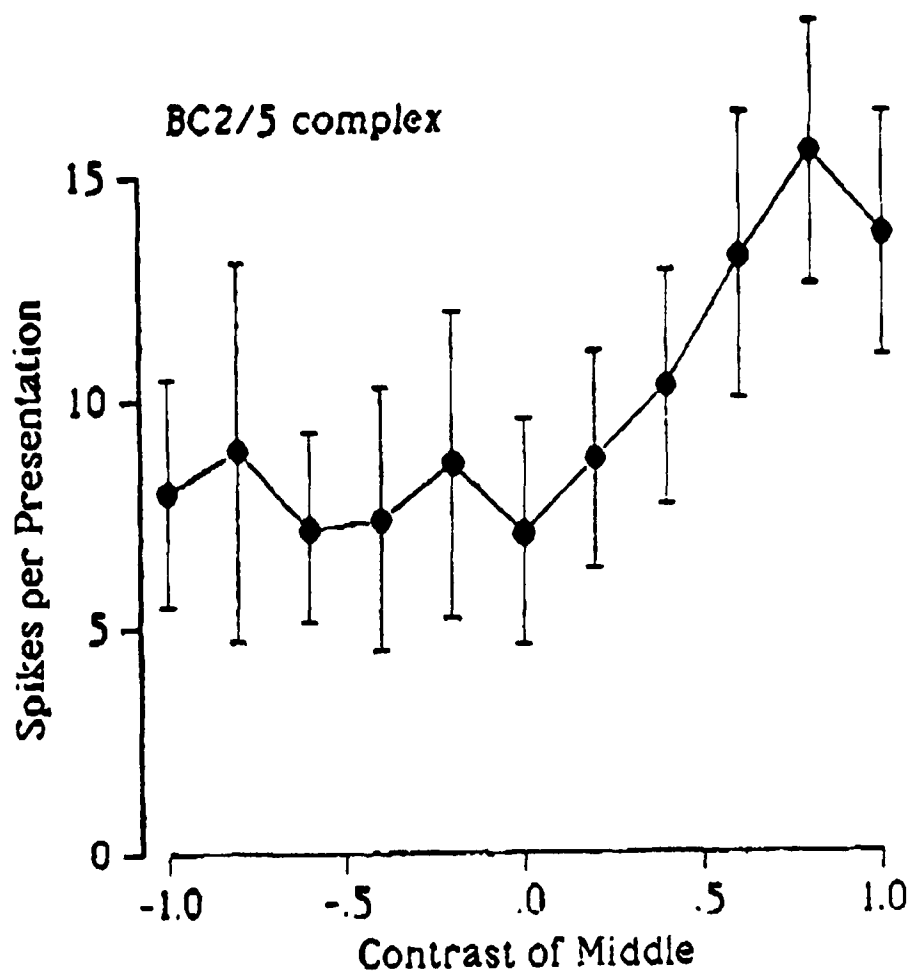


FIG. 1

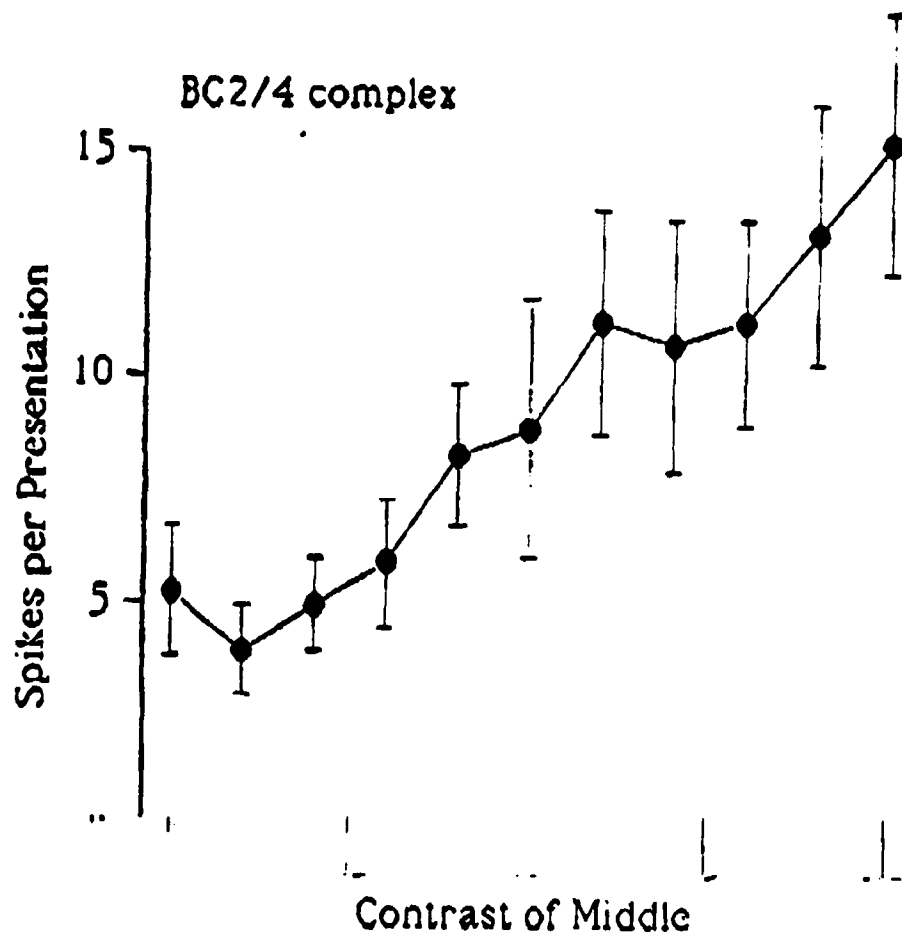


FIG. 2

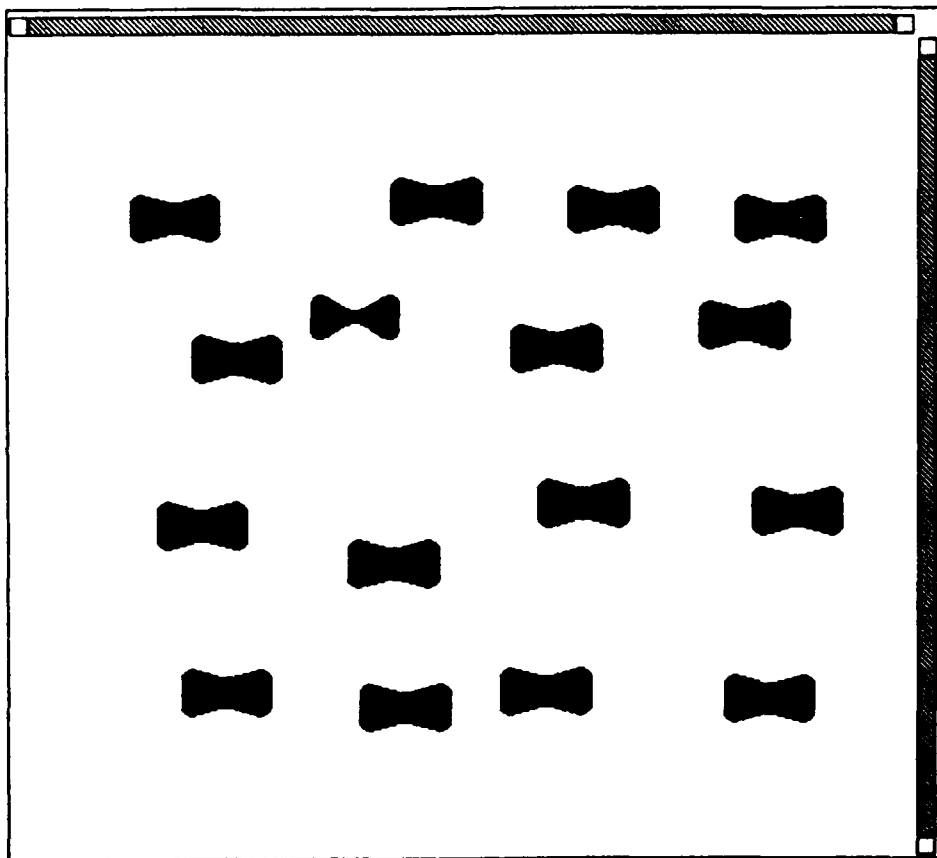


FIG. 3

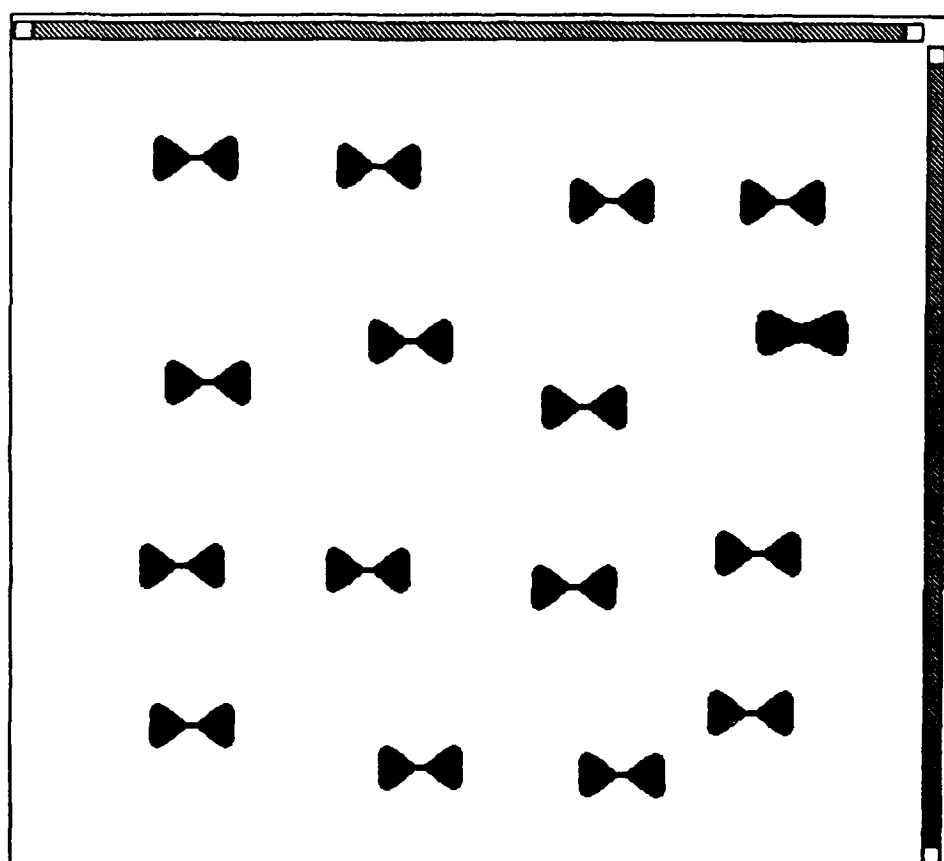


FIG. 3 (cont.)

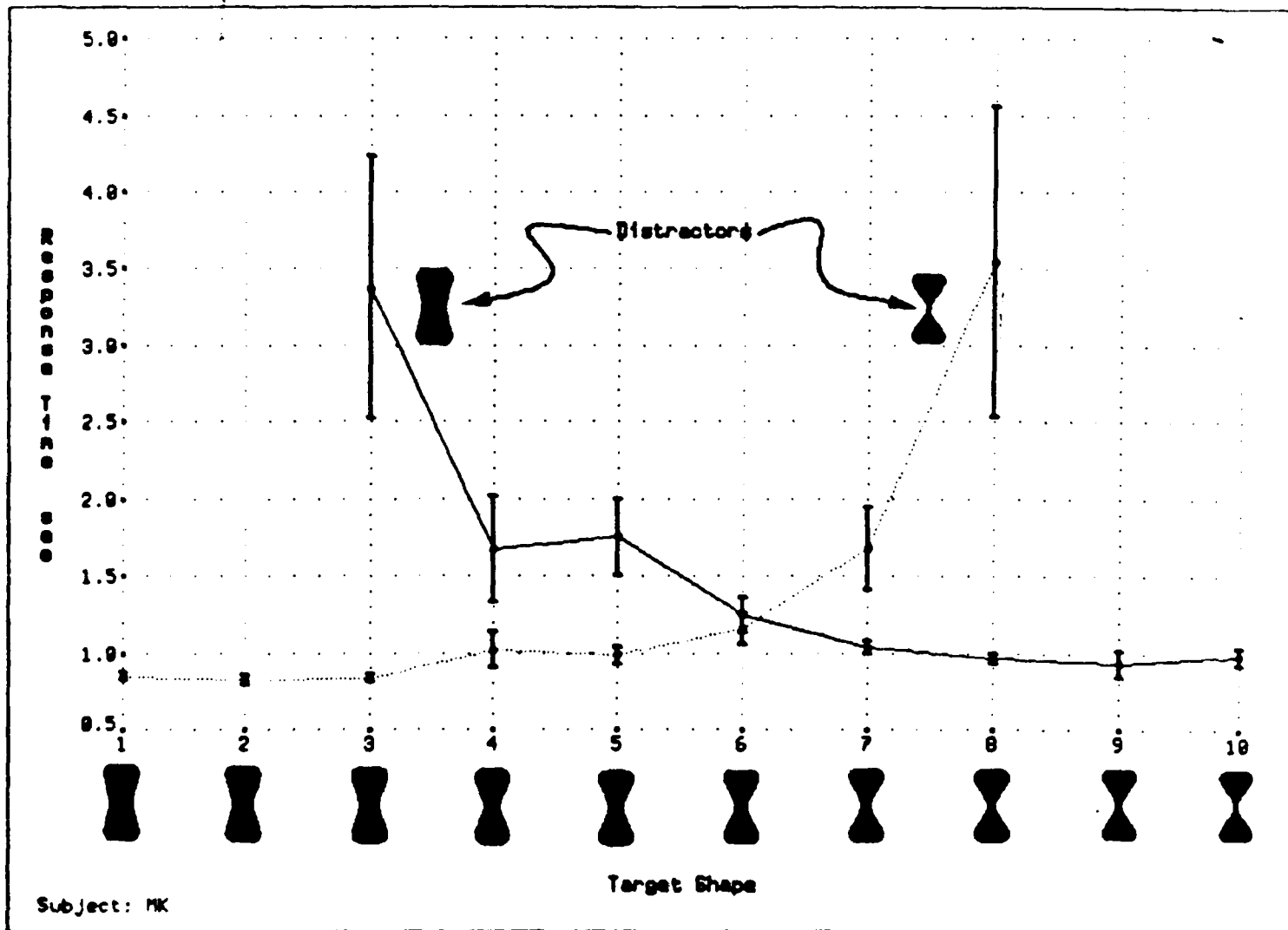


FIG. 3 (cont.)

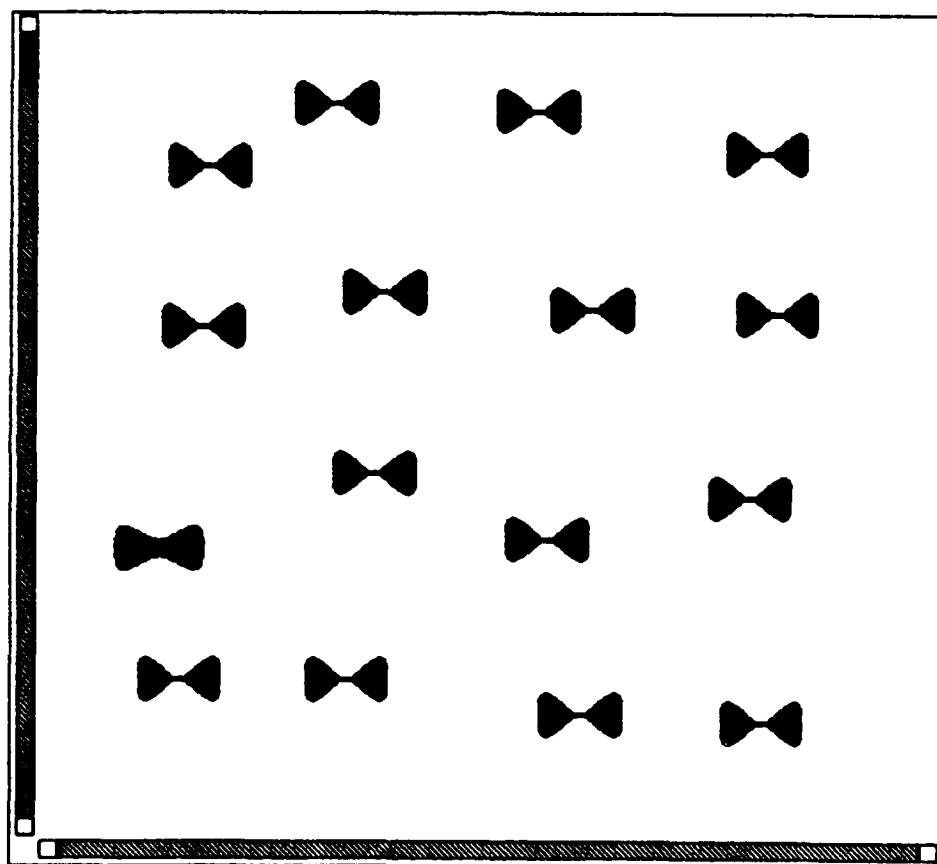


FIG. 3 (cont.)

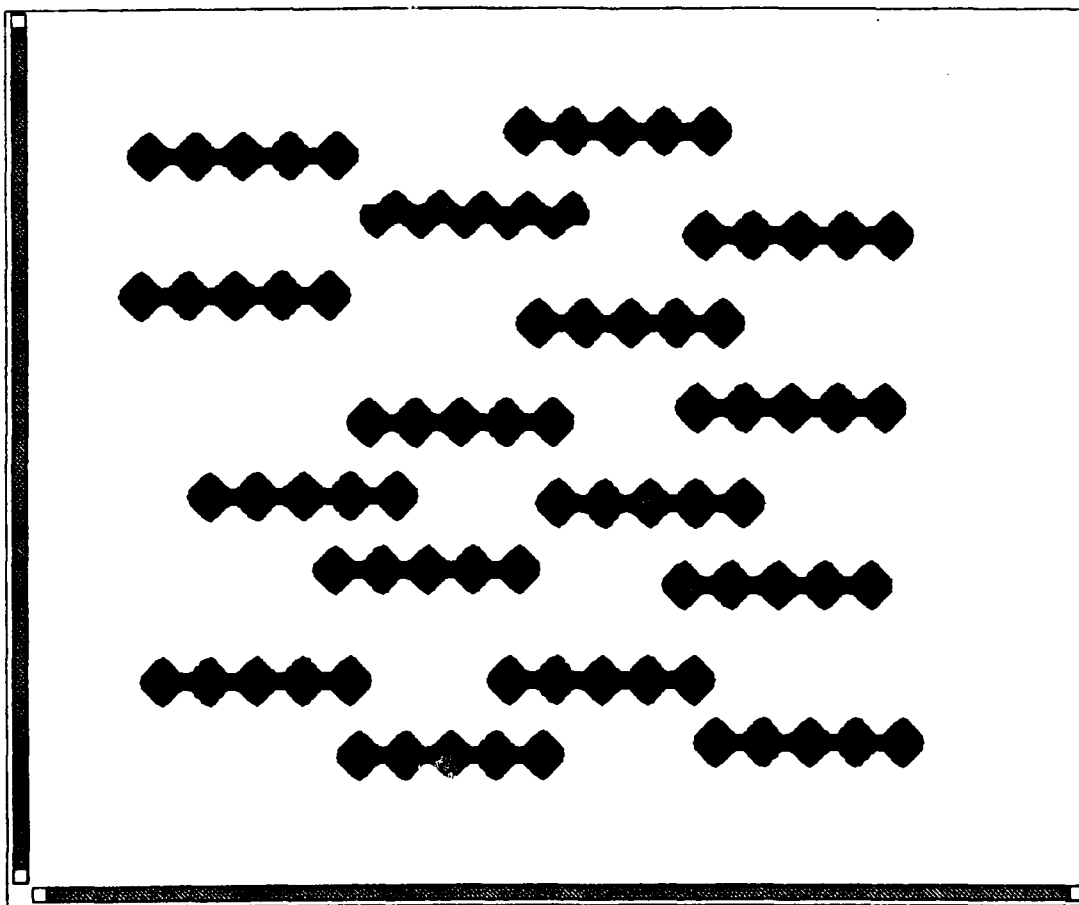


FIG. 4 (cont.)

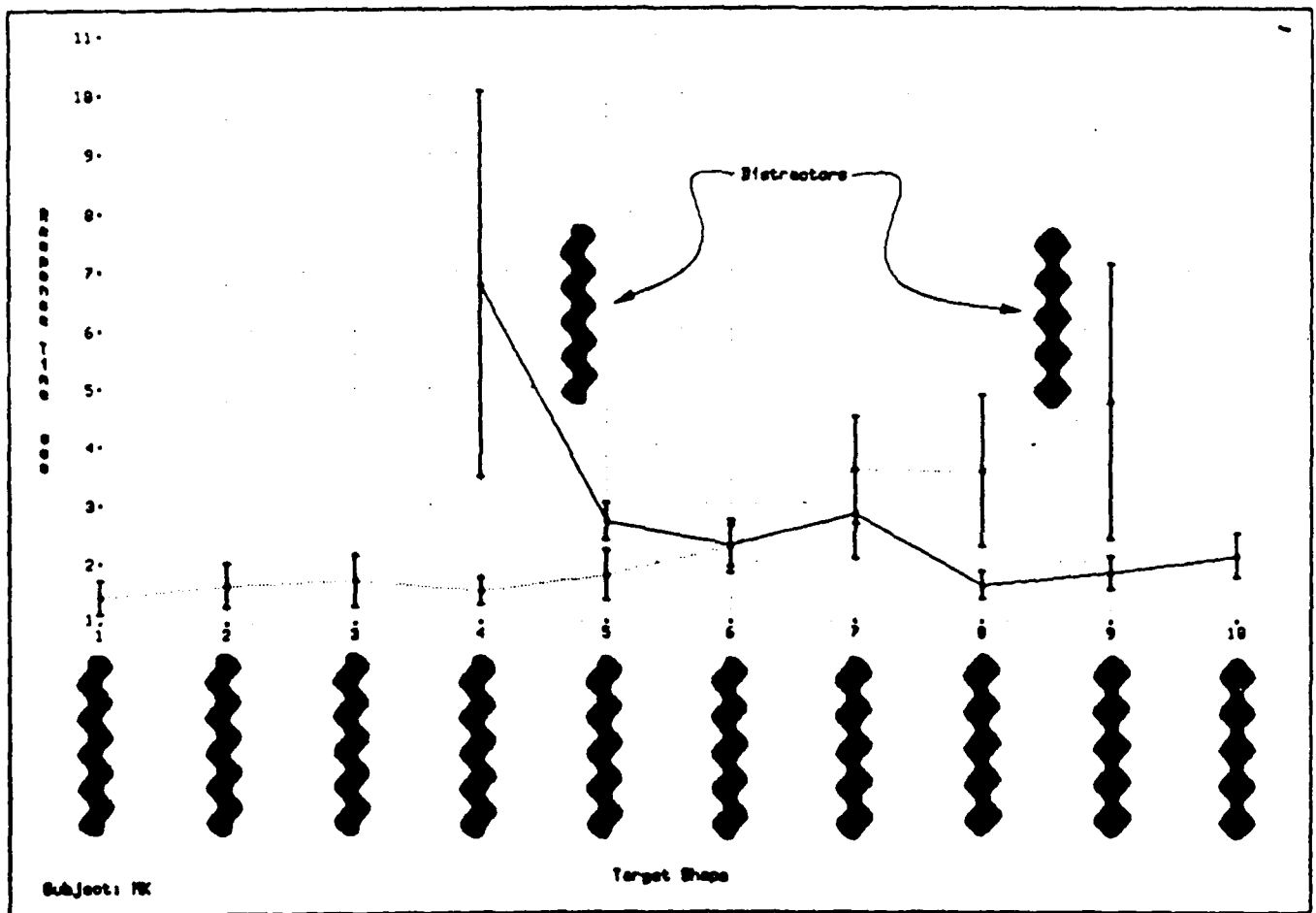


FIG. 4 (cont.)

Appendix I

Logical/Linear Operators for Measuring Orientation and Curvature, Technical Report TR-CIM-90-06, McGill Research Center for Intelligent Machines, Montreal.

McGill Research Centre for Intelligent Machines



McGill University



**Logical/Linear Operators for Measuring
Orientation and Curvature**

Lee A. Iverson* Steven W. Zucker*†

TR-CIM-90-06 July, 1990

Computer Vision and Robotics Laboratory
McGill Research Centre for Intelligent Machines
McGill University
Montreal, Quebec, Canada

* McGill University, Dept. of Electrical Engineering

† Fellow, Canadian Institute for Advanced Research

Postal Address: 3480 University Street, Montreal, Quebec, Canada H3A 2A7
Telephone: (514) 398-6319 Telex: 05 268510 FAX: (514) 283-7897
Network Address: mrcim@larry.mrcim.mcgill.edu

Logical/Linear Operators for Measuring Orientation and Curvature

Lee A. Iverson

Steven W. Zucker

Abstract

We propose a language which unifies aspects of linear operator theory and boolean logic to define *logical/linear* (L/L) operators. A family of these operators appropriate for measuring the low-order differential structure of image curves is developed. These L/L operators are derived by decomposing a linear model into logical components to ensure that certain logical preconditions for the existence of an image curve are upheld. The resulting operators allow for coarse measurement of curvilinear differential structure (orientation and curvature) while successfully segregating edge- and line-like features. By thus reducing the incidence of false-positive responses, these operators are a substantial improvement over (thresholded) linear operators which attempt to resolve the same class of features. Furthermore, these logical decomposition techniques naturally overcome the need for the *ad hoc* post-processing linear methods normally rely on. Finally, evidence is advanced which would indicate that these operators are appropriate models for end-stopped and non-end-stopped simple cells in primary visual cortex.

Index items: Edge detection, feature extraction, image processing, computer vision, non-linear operators.

Acknowledgements

We thank Allan Dobbins for his many contributions and insights.

Research supported by grants from the AFOSR, MRC and NSERC.

Contents

1	Introduction	1
1.1	Image Curves	2
2	A Logical/Linear Framework for Line Operators	4
2.1	Logical/Linear Decomposition	4
3	Designing L/L Operators for Image Curves	7
3.1	Testing Lateral Extrema	7
3.1.1	Other Curve Types	10
3.2	Testing Tangential Continuity	11
3.3	Curvature and End-Stopping	16
4	Results	18
5	Discussion	27
5.1	Biological Implications	27
5.2	Neural Networks	28
5.3	Conclusions	28
A	Logical/Linear Combinators	30
	Bibliography	33

List of Figures

1	Possible image curve configurations	2
2	Image curve	3
3	Cross sections of image lines and edges.	8
4	Approximation of G' by two G operators.	8
5	Responses of L/L and linear operators near step.	10
6	Responses of operators near end of line.	12
7	Schematic of the half-field decomposition and line endings.	12
8	Splitting of 1D gaussian operator into components	14
9	Curvature responses of simple cells	17
10	Response profiles of a family of curvature-selective operators.	19
11	Artificial test image	19
12	Family of responses to the artificial test image.	20
13	The effect of "soft" L/L operators on test image responses	21
14	The effect of added noise on test image responses	22
15	L/L responses to a fingerprint image.	23
16	L/L responses to a remote sensing image of logging roads.	23
17	L/L responses to a cerebral angiogram.	24
18	Comparison of edge responses for a noisy test figure and (a) the L/L edge operators, (b) Canny's operator.	25
19	Graphs of L/L combinators varying ρ	31
20	Network implementation of L/L AND	32

1. Introduction

There is no shortage of so-called "edge-detectors" and "line-detectors" in computer vision. These operators are intended to respond to lines and edges in intensity images. Many different designs have been proposed, based on a range of optimality criteria (e.g. [Heuckel, 1971; Canny, 1986]) and, in the end, many of these designs exhibit properties in common with biological vision systems [Jones and Palmer, 1987]. While this agreement between mathematics and physiology is encouraging, there is still dissatisfaction with these operators—despite their 'optimal' design, they do not work sufficiently well to support subsequent analysis. Part of the problem is undoubtedly the myopic perspective to which such operators are restricted, suggesting the need for more global interactions [Zucker *et al.*, 1988]. But we believe that more can be done locally, and that another significant part of the problem stems from the types of models on which the operators are based and the related mathematical tools that have been invoked to derive them. In this paper we introduce an approach to operator design that differs significantly from the standard practice, and illustrate how it can be used to design non-linear operators for locating lines and edges. Finally, relating these non-linearities back to physiology suggests explanations for some of the characteristics of simple cells left unexplained by linear models.

A standard model used in the design of edge operators involves two components: an ideal step edge plus additive gaussian noise. This model was proposed in one of the first edge detector designs [Herskovitz and Binford, 1970], and has continued through the most recent [Canny, 1986; Deriche, 1987]. Thus it is no surprise that the solution resembles the product of two (operator) terms, one to smooth the noise (e.g. a gaussian) and the other to locate the edge (e.g. a derivative or a step function).

While some of the limitations of the ideal step edge model have been addressed elsewhere (e.g. [Horn, 1977; Leclerc and Zucker, 1984]), a perhaps more important limitation of the operator design has not been considered. It is assumed that in viewing a small local region of the image, only a single section of one edge is being examined. This may be a valid simplification in some continuous limit, but it is definitely not valid in digital images. Many of the systematic problems with edge and line detectors occur when structure changes within the local support of the operator (e.g. several edges or lines coincide). Since these singularities are not dealt with by the noise component of the model either, the linear operator behaves poorly in their vicinity.

In particular, curve-detecting operators are usually designed to respond if a certain intensity configuration exists locally (see Figure 1(a)). A signal estimation component of the operator is then incorporated in the design to filter local noise (Figure 1(b)). However, significant contrast changes are rarely noise—they are more likely to be the result of a set of distinct objects whose images project to coincident image positions (Figure 1(c)). An operator which claims to 'detect' or 'select' a certain class of image features must continue to do so in the presence of such confounding information.

We propose that image operators should be designed to respond positively to the expected image structure, and *to not respond at all* when such structure is *not* present. Simple linear operators achieve the first of these goals, but in order to fulfill the second we must incorporate a more direct verification of the logical preconditions for a given feature into the operator itself. We accomplish this by decomposing the linear operator into components which correspond to the logical assertions of the preconditions. When the expected image structure is present, a *boolean* combination of these responses produce a *linear* response, whereas if any of the logical conditions

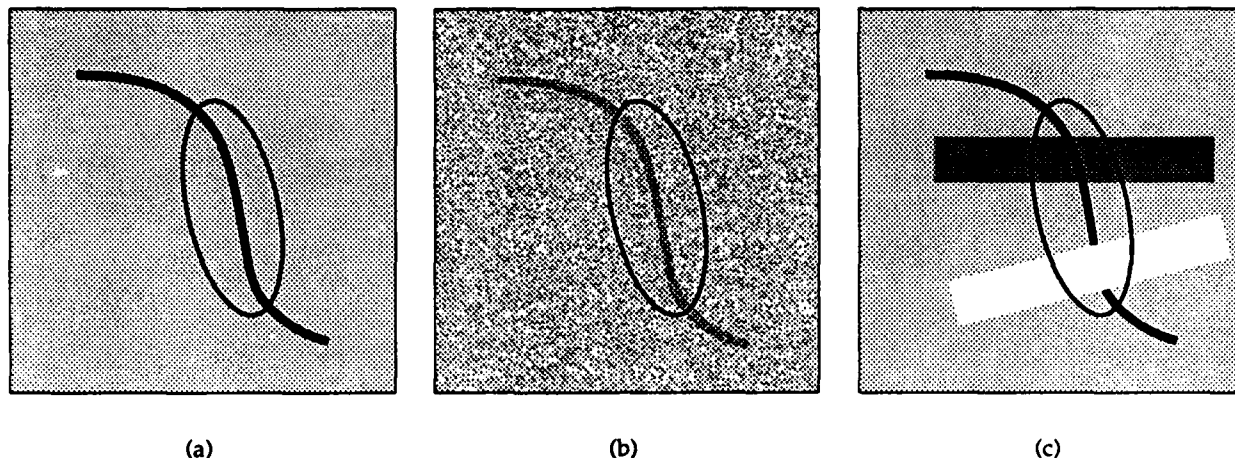


Figure 1: A set of potential image curve configurations which must be considered in the design of operators. An ideal image (a) of a black curve on a white background; a noisy image (b) of a lower-contrast version of the same curve; an obscured version (c) of the ideal image. The oval in each image represents the spatial support of a local operator.

are violated the response will be suppressed non-linearly. Because these operators unite elements from boolean logic and linear operator theory, we refer to them as *logical/linear* (L/L) operators.

1.1 Image Curves

For consistency we shall adopt the following terminology. *Edges* are the curves which separate lighter and darker areas of the image—the perceived discontinuities in the intensity surface; *Lines* are those curves which might have been drawn by a pen or pencil (sometimes referred to as *bars* in other work [Marr, 1982]). *Image curves* may be either of these. Two independent properties describe these image curves: their structures along the tangential and the normal directions. Tangentially, lines and edges are similar; it is the cross-sectional structure that differentiates them.

Formally, let $I: \mathbb{R}^2 \rightarrow \mathbb{R}$ be an analytic intensity surface (an image) and $\alpha: S = (s_0, s_1) \rightarrow \mathbb{R}^2$ a smooth curve parametrized by arc length (see Figure 2). The orientation $\theta(s)$ is the direction of the tangent $\tau(s)$, a unit vector in the direction of $\alpha'(s)$, and the normal vector $n(s)$ is a unit vector in the direction $\alpha''(s)$. Define the cross-section β_s of α

$$\beta_s(t) = I(\alpha(s) + t n(s)), \quad s \in S, t \in \mathbb{R}. \quad (1)$$

The three kinds of image curve are then defined:

DEFINITION 1: α is a **Positive Contrast Line** iff

$$\exists \epsilon > 0: \beta_s(0) = \max_{-\epsilon < \delta < \epsilon} \beta_s(\delta), \quad s \in S. \quad (2a)$$

DEFINITION 2: α is a **Negative Contrast Line** iff

$$\exists \epsilon > 0: \beta_s(0) = \min_{-\epsilon < \delta < \epsilon} \beta_s(\delta), \quad s \in S. \quad (2b)$$

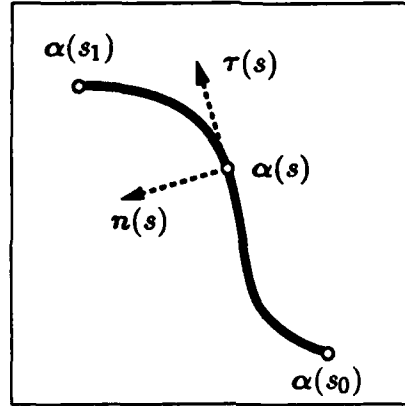


Figure 2: An image curve $\alpha: S = (s_0, s_1) \rightarrow \mathbb{R}^2$ with the tangent $\tau(s)$ and normal $n(s)$.

DEFINITION 3: α is an **Edge*** iff

$$\beta'_s > 0 \text{ and } \exists \epsilon > 0: |\beta'_s(0)| = \max_{-\epsilon < \delta < \epsilon} |\beta'_s(\delta)|, \quad s \in S. \quad (2c)$$

These are the structural preconditions that define image curves: they *must* hold over their length (S) and width (ϵ). Linear operators do respond when these conditions are met. However, they also respond in characteristic situations in which the conditions are clearly *not* met. These responses are referred to as *false-positives*. The current analysis will focus on a mechanism for avoiding three kinds of false-positives typical of linear operators:

1. *Confusion between lines and edges*: Lines and edges are differentiated in terms of their cross-sections. For accurate identification the logical conditions on the cross-section must be satisfied, and in each case we will show that a linear operator tests them incompletely.
2. *Merging or interference between nearby curves*: The local continuity of image curves is important for resolving and separating nearby features. Linear operators interfere with testing continuity by filling in gaps between nearby features.
3. *Smoothing out discontinuities or failing to localize line-endings*: The locations of the discontinuities and end-points of a curve are fundamental to higher-level descriptions [Koenderink and van Doorn, 1976; 1982] and [Koenderink, 1984]. Linear operators systematically interfere with their localization by arbitrary smoothing.

*Note that the definition of a line is sign specific, whereas an edge is defined as a locus of local maxima or minima. There is no conceptual difference between edges that go from light to dark moving towards their centre of curvature and those of the opposite contrast; they are both edges.

2. A Logical/Linear Framework for Line Operators

The three qualitatively different kinds of image curves defined in §1.1 imply three distinct sets of preconditions for the existence of an image curve. As noted previously, the curve description process must respect these distinctions. Focusing for the moment on the line condition of Definition 1*, we begin by adopting an oriented, linear line operator similar to the one described in [Canny, 1986].

Canny adopted the assumption of linearity to facilitate noise sensitivity analysis, and relied on post-processing to guarantee locality and selectivity of response. He arrived at a line operator whose cross-section is similar to a gaussian second-derivative, and an edge operator similar to a gaussian first derivative. Neurophysiologists [Movshon *et al.*, 1978; Schumer and Movshon, 1984; Jones and Palmer, 1987] and psychophysicists [Shapley and Lennie, 1985] have adopted such linear models to capture many of the functional properties of the early visual system, and they are relatively easy to analyse mathematically (e.g. using Fourier analysis). These models are also attractive from a strictly computational point of view because they exhibit most of the properties required of a measurement operator for image curves. However, they also exhibit the false-positive responses described above (partially shown in Figures 5 and 6).

To limit these false-positives, we relax the assumption of linearity and test the necessary local conditions explicitly. This is accomplished by decomposing the operator into components which measure the conditions to be tested. The resulting operator will appear to be linear as long as all of these conditions are fulfilled, and thus for many of the standard 'probe stimuli' used by visual neurophysiologists. We will refer to these operators as logical/linear (L/L) operators.

2.1 Logical/Linear Decomposition

Consider the necessary conditions for a decomposition which will be specific to the design preconditions but still retain some of the desirable properties of a linear operator.

DEFINITION 4: *A non-linear operator is said to be a Logical/Linear operator for a given feature iff*

1. *It consists of a linear operator which is decomposed into a set of linear component operators whose sum is identical to the initial operator. These components correspond to measurements of the logical conditions for the feature's existence.*
 2. *The overall operator response is positive only if the logical conditions for the desired feature are fulfilled. When this is so, the operator should act identically to the original linear operator.*
-

Thus, given a linear operator which is conditionally optimal (e.g. in terms of noise immunity and localization) a linear/logical decomposition of this should retain such optimality as long as the conditions on the component responses hold.

The first condition above suggests that the required decomposition is a partition:

*That this analysis is applicable to the other cases will be shown in §3.1.1

DEFINITION 5: For a function $f: X \rightarrow \mathbb{R}$, a set of functions $\Pi = \{\pi_1, \pi_2, \dots\}$ where $\pi_i: X \rightarrow \mathbb{R}$ is a partition of f iff for all $x \in X$

$$\sum_{\pi_i \in \Pi} \pi_i(x) = f(x). \quad (3)$$

Many such partitions are clearly possible; we must choose the components Π such that they embody the required logical features.

The combinators to fulfill the second condition of Definition 4 can be derived from a mapping of the real line to logical values. Assume that positive operator responses represent confirmation of a logical condition (logical *true*) and that negative responses represent rejection (logical *false*). To derive the numeric \Rightarrow logical mapping, we adopt the principle that all confirming evidence should be combined if the logical condition holds, and contradictory evidence combined if the condition fails. This leads to the following set of logical/linear combinators:

DEFINITION 6: The basic set of Logical/Linear Combinators Δ , ∇ , absolute value, and ϕ (rectification) are given by

$$x \Delta y \triangleq \begin{cases} x + y, & \text{if } x > 0 \wedge y > 0; \\ y, & \text{if } x > 0 \wedge y \leq 0; \\ x, & \text{if } x \leq 0 \wedge y > 0; \\ x + y, & \text{if } x \leq 0 \wedge y \leq 0. \end{cases} \quad x \nabla y \triangleq \begin{cases} x + y, & \text{if } x > 0 \wedge y > 0; \\ x, & \text{if } x > 0 \wedge y \leq 0; \\ y, & \text{if } x \leq 0 \wedge y > 0; \\ x + y, & \text{if } x \leq 0 \wedge y \leq 0. \end{cases} \quad (4a)$$

$$|x| \triangleq \begin{cases} x, & \text{if } x > 0; \\ -x, & \text{if } x \leq 0. \end{cases} \quad \phi(x) \triangleq \begin{cases} x, & \text{if } x > 0; \\ 0, & \text{if } x \leq 0. \end{cases} \quad (4b)$$

Observe that $|x| = x \Delta -x$ and $\phi(x) = -(0 \Delta x)$.

Using these combinators, define a generative syntax for L/L expressions.

DEFINITION 7: A Logical/Linear Expression of the real variables $v = \{v_1, \dots, v_n\}$ is any expression e in the language E defined by the following grammar (for the a 's being real constants):

$$\begin{aligned} E &\rightarrow v_i & E &\rightarrow a E \\ E &\rightarrow E \Delta E & E &\rightarrow E \nabla E \\ E &\rightarrow |E| & E &\rightarrow \phi(E) \end{aligned}$$

Consider the universe of vectors U in \mathbb{R}^n excluding the axes*

$$U = \{v \in \mathbb{R}^n \mid v_i \neq 0, 1 \leq i \leq n\} \quad (5)$$

and the subspaces $\{e\}_+ = \{v \in U \mid e > 0\}$.

*For real-valued variables, the exclusion of the axes needed to demonstrate logical equivalence is not problematic because of the singularity of this restriction.

THEOREM 1: (LOGICAL) For the language of L/L expressions E , the set of all sets $\{e\}_+$ and their complements $\overline{\{e\}}_+ = U - \{e\}_+$ (for $e \in E$) form a Boolean algebra with meet Δ , join ∇ and complement $-$.

The following equivalences can be derived directly from Definition 6, for all $e_1, e_2 \in E$:

$$\{-e_1\}_+ = \overline{\{e_1\}}_+, \quad (6a)$$

$$\{e_1 \Delta e_2\}_+ = \{e_1\}_+ \cap \{e_2\}_+, \quad (6b)$$

$$\{e_1 \nabla e_2\}_+ = \{e_1\}_+ \cup \{e_2\}_+. \quad (6c)$$

It is easy to verify that these sets form a field with the help of the equivalences above (e.g. the equivalence of Δ and \cap ensures that if X and Y are members then $X \Delta Y$ is also). Furthermore, these meet, join and complement operators are clearly isomorphic with the standard set-theoretic \cap , \cup and complement. The further observation that \emptyset and U are the bounds of this field ensure that this system is a Boolean algebra. ([Sikorski, 1960], p. 3) ■

The following equivalences can also be derived directly and are useful.

$$\begin{aligned} \{ae\}_+ &= \{e\}_+ \text{ if } a > 0 & \{ae\}_+ &= \overline{\{e\}}_+ \text{ if } a \leq 0 \\ \{|e|\}_+ &= U & \{\phi(e)\}_+ &= \{e\}_+. \end{aligned}$$

To expose the linearity inherent in these operators, it is necessary to consider the 2^n minimal polynomials

$$P_j = q_1 \Delta q_2 \Delta \cdots \Delta q_n \quad (7)$$

where $q_i = v_i$ if bit i in the binary representation of j is zero, $q_i = -v_i$ if bit i is one.

THEOREM 2: (LINEAR) Any L/L expression e is linear within the subspace $\{P_j\}_+$ of any minimal polynomial P_j .

Any Boolean polynomial can be equivalently expressed as the join of minimal polynomials or the lower bound \emptyset ([Birkhoff and MacLane, 1977], p. 370). Thus $\{e\}_+$ can be expressed as the ∇ of a group of such minimal polynomials (the *disjunctive canonical form* (DCF) of e). Without loss of generality, consider a particular such polynomial P_j . Noting that every element v_i in a vector $v \in \{P_j\}_+$ has a fixed sign, Definition 6 guarantees that Δ is linear on the subspace defined by $\{P_j\}_+$ (for fixed sign arguments, the branch chosen in the Δ is fixed). Thus, any minimal polynomial P_j is linear on $\{P_j\}_+$.

Consider now the DCF of e . We know that each P_j in this DCF is both linear and of constant sign on $\{P_j\}_+$. By the same reasoning as for Δ above, we can state that ∇ is linear if its arguments have constant sign, and thus the DCF of e is a linear combination of expressions which are guaranteed linear on $\{P_j\}_+$. Therefore, e is also linear on every $\{P_j\}_+$. ■

Thus, the proposed combinators satisfy the goals of Definition 4, and have both logical and linear properties. We will rely on both of these properties in constructing the L/L image operators. Refer to Appendix A for a more detailed analysis of the implications of these L/L constructs.

3. Designing L/L Operators for Image Curves

3.1 Testing Lateral Extrema

We begin by localizing the cross-section of a *positive contrast line* (Definition 1). A necessary condition for the existence of such a line is a local extremum in intensity (Figure 3 is a display of typical 1D cross-sections of these intensity phenomena), we will first consider the operator structure perpendicular to its preferred direction. This specializes to the prior problem of locating extrema in the cross-section β_s .

A local extremum in a one-dimensional signal* $\beta(x)$ exists only at those points where

$$\beta'(x) = 0 \text{ and } \beta''(x) \neq 0. \quad (8)$$

Estimating the location of such zeroes in the presence of noise is normally achieved by locating zero-crossings, thus in practice these conditions become

$$\beta'(x - \epsilon) > 0 \text{ and } \beta'(x + \epsilon) < 0 \text{ and } \beta''(x) < 0 \quad (9)$$

An operator which can reliably restrict its responses to only those points where these conditions hold will only respond to local maxima in a one-dimensional signal.

A set of noise-insensitive linear derivative operators (or 'fuzzy derivatives' [Koenderink and van Doorn, 1987]) are the various derivatives of a Gaussian envelope.

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}, \quad (10a)$$

$$G'_\sigma(x) = -\frac{x}{\sigma^2} G_\sigma(x), \quad (10b)$$

$$G''_\sigma(x) = \frac{x^2 - \sigma^2}{\sigma^4} G_\sigma(x). \quad (10c)$$

When convolved over a one-dimensional signal these give noise-insensitive estimates of the derivatives of the signal. These smooth derivatives of the signal are

$$\beta'_\sigma(x) = \beta'(x) * G_\sigma(x) = \beta(x) * G'_\sigma(x), \quad (11a)$$

$$\beta''_\sigma(x) = \beta''(x) * G_\sigma(x) = \beta(x) * G''_\sigma(x). \quad (11b)$$

THEOREM 3: $\beta'_\sigma(x - \epsilon) > 0$ and $\beta'_\sigma(x + \epsilon) < 0$ and $\beta''_\sigma(x) < 0$ are sufficient conditions for a local maximum in the signal $\beta(x)$.

The identities of (11) show that these conditions are necessary and sufficient for the existence of a local maximum in $\beta_\sigma(x) = \beta(x) * G_\sigma(x)$. The maximum principle for the heat equation ([Protter and Weinberger, 1984], p. 161) implies that convolution by a gaussian cannot introduce new maxima. Thus the above conditions imply the existence of a maximum in $\beta(x)$. ■

* Assume $\beta(x)$ is sufficiently differentiable.

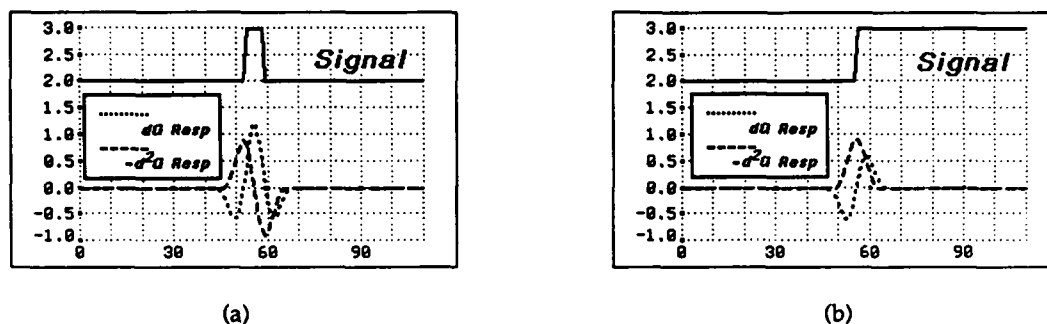


Figure 3: Cross sections of image lines and edges. A line in an intensity image (a) is located at the peak of its cross-section. Note that this coincides with a zero in the derivative β' and a negative second derivative β'' . An intensity edge (b) occurs at peaks in the derivative β' of the cross-section. The derivatives shown are derived from convolution by G' and G'' operators with $\sigma = 3$.

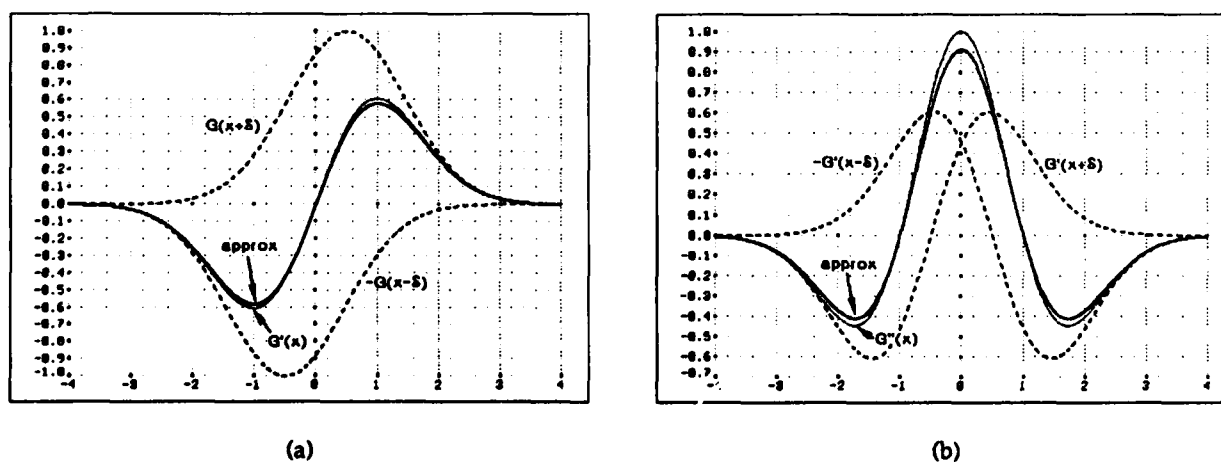


Figure 4: Approximation of (a) G' by the difference of two G operators, and (b) G'' by the difference of two G' operators.

This suggests a practical method for locating maxima in a noisy one-dimensional signal. Comparing the results of convolutions by derivatives of gaussians will allow us to determine the points where Theorem 3 holds. The loci of such points will form distinct intervals with widths $\leq 2\epsilon$. The parameter σ determines the amount of smoothing used to reduce noise-sensitivity.

Using central limits, observe that:

$$f'(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} (f(x + \epsilon) - f(x - \epsilon)). \quad (12)$$

Thus for the fuzzy derivatives of β , one would expect to find that

$$\beta''_{\sigma}(x) \approx \frac{1}{2\epsilon} (\beta'_{\sigma}(x + \epsilon) - \beta'_{\sigma}(x - \epsilon)) \quad (13)$$

with the quality of the approximation a function of ϵ . Thus the conditions in Theorem 3 can be derived from the derivative $\beta'_{\sigma}(x)$ ---a linear combination of two points will give $\beta''_{\sigma}(x)$. More specifically, we adopt the approximation $G''(x) \approx (G'_{\sigma}(x + \epsilon) - G'_{\sigma}(x - \epsilon))/2\epsilon$, where $\epsilon/\sigma \leq 1$. Figure 4(b) shows that for $\epsilon = \sigma/2$ this is an excellent approximation.

Thus, convolution by G' allows testing of all three conditions in Theorem 3 simultaneously. Using the L/L combinators of §2.1 we are now able to define a one-dimensional operator which has a positive response only within a small range of local maxima.

DEFINITION 8: The One-Dimensional Maxima Operator Ψ_{σ} is given by

$$\Psi_{\sigma}(\beta)|_x = (\Psi_{\sigma}^{+}(x) * \beta(x)) \Delta (\Psi_{\sigma}^{-}(x) * \beta(x)) \quad (14a)$$

where

$$\Psi_{\sigma}^{-} = -G'_{\sigma}(x + \epsilon)/2\epsilon, \quad (14b)$$

$$\Psi_{\sigma}^{+} = G'_{\sigma}(x - \epsilon)/2\epsilon. \quad (14c)$$

We rely on a shorthand notation when describing operators as combinations of linear or logical/linear components which are evaluated at all x ; (14a) can be equivalently expressed as

$$\Psi_{\sigma} = \Psi^{-} \Delta \Psi^{+}. \quad (15)$$

Clearly then the key advantage of this Ψ_{σ} operator is that:

LEMMA 4: The response $\Psi_{\sigma}(\beta)|_x$ will be positive only if there is local maximum in β_{σ} within the region $[x - b\sigma, x + b\sigma]$.

By Theorem 1 we know that $\Psi_{\sigma}(\beta)|_x > 0$ implies that both (16b) and (16a) must also be positive. Equations 11 then imply that

$$\Psi_{\sigma}^{-}(x) * \beta(x) = -\beta'_{\sigma}(x + \epsilon)/2\epsilon \quad (16a)$$

$$\Psi_{\sigma}^{+}(x) * \beta(x) = \beta'_{\sigma}(x - \epsilon)/2\epsilon \quad (16b)$$

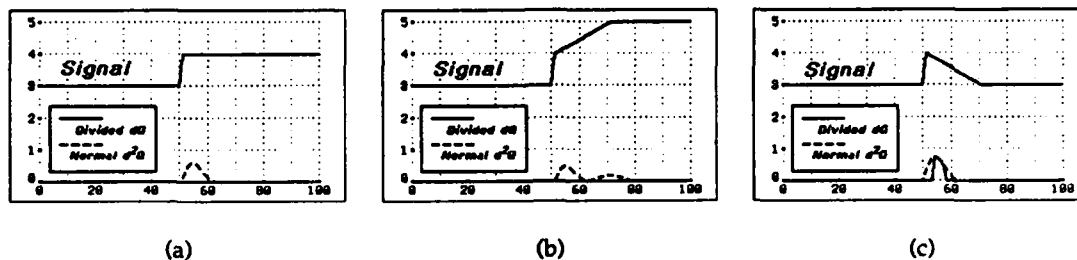


Figure 5: Responses of L/L and linear maxima operators near step. Rectified responses of the G'' operator and divided G' operators are shown for three step edges with varying slopes of the upper region. It can be seen that the non-linear operator blocks the unwanted response near a step which is not also a local maximum (a & b), but that when it is a local maximum (c) it does respond. The linear G'' operator, however, responds in each of these cases, exhibiting consistent (and erroneous) displacement of the peak response.

Thus a positive response requires that $\beta'_\sigma(x - \epsilon) > 0$ and $\beta'_\sigma(x + \epsilon) < 0$, which in turn imply the presence of a local maximum on β_σ between $x - \epsilon$ and $x + \epsilon$.*

The performance improvement from introducing this non-linearity is considerable. The linear operator exhibits consistent patterns of false positive responses. The simplest example is the response near a step (see Figure 5). The linear operator displays a characteristic (false) peak response when the step is centered over one of the zeroes in the operator profile. The logical/linear Δ operation prevents this error since both G' halves of the operator register derivatives in the same direction and so do not fulfill the conditions of (9). The L/L operator will respond positively only in the case that the slope above the step is negative (i.e. only when the transition point is also a local maximum).

3.1.1 Other Curve Types

The extension of this analysis to the other curve types in §1.1 is straightforward. The most obvious observation is that the above analysis can be repeated entirely with a simple change of sign so as to be specific for one-dimensional minima, which correspond to negative contrast lines.

Slightly more complicated is the case for edges, where two related conditions must be combined (Definition 3). For this case, the following conditions (analogous to equation (8)) must be verified

$$\beta'(x) > 0 \text{ and } \beta''(x) = 0 \text{ and } \beta'''(x) < 0 \quad (17a)$$

or

$$\beta'(x) > 0 \text{ and } \beta''(x - \epsilon) > 0 \text{ and } \beta''(x + \epsilon) < 0 \text{ and } \beta'''(x) < 0 \quad (17b)$$

*Observe that although the local maximum in β_σ is guaranteed to fall within this region, the corresponding maximum in β is not necessarily so restricted. Qualitatively however, we can rely on the observation that the maxima for a signal will converge on the centroid of that signal under heat propagation (or as we convolve with larger and larger gaussians). Considering the features of β in isolation then, we can state that the smoothing will cause the location of the local maximum in β_σ to shift towards the centroid of the local intensity distribution.

Mirroring the analysis above, the verification of these conditions can be realized in an L/L operator by defining the components

$$\Psi^* = G'_\sigma(x) \quad (18a)$$

$$\Psi_2^- = G''_\sigma(x - \epsilon) \quad (18b)$$

$$\Psi_2^+ = -G''_\sigma(x + \epsilon) \quad (18c)$$

and the combination

$$\Psi = \Psi^* \Delta \Psi_2^- \Delta \Psi_2^+ \quad (18d)$$

or, equivalently,

$$\Psi = \Psi^* \Delta (\Psi^* - \Psi^-) \Delta (\Psi^* - \Psi^+) \quad (19)$$

where Ψ^- and Ψ^+ are as defined in (14).

Supporting this use of the third derivative as a measure of edge confidence is the observation that a blurred step edge has vanishing even derivatives and positive odd derivatives. The description of an edge adopted in (17) is clearly consistent with this observation. It is also more selective than the 'zero-crossing of a second-derivative,' [Marr and Hildreth, 1980] which is only one of the logical preconditions which this operator depends on for a response. One can therefore expect less of a problem of non-edge signals generating edge-like responses with this operator than with those less specific.

3.2 Testing Tangential Continuity

So far only the cross-sectional (normal) image structure (β_s) has been discussed. In order to extend this result to two-dimensions, we must examine the tangential (curvilinear) structure of the curves (α). By (2) we must verify the local connectedness of candidate curves. In addition, the extraction of orientation-specific measures was deemed essential for further processing. In this section, these problems will be addressed by imposing a further tangential structure on the operator.

We will follow the same course as for the cross-section—first a linear structure is proposed which will be decomposed to reveal linear measurement operators for the components of the logical preconditions. Assume initially that the intensity variation along a curve is everywhere smooth and corrupted only by additive gaussian noise. This suggests that the response should be noise-filtered with a linear gaussian operator. Incorporating the prior conditions of §3.1, we arrive at the first two-dimensional candidate operator:

DEFINITION 9: A preliminary Two-Dimensional Positive Contrast Line Operator Ψ is given by $\Psi = \Psi^- \Delta \Psi^+$ where the lateral components Ψ^- and Ψ^+ are given by

$$\Psi^-(x, y) = -G'_{\sigma_y}(y + \epsilon) G_{\sigma_x}(x), \quad (20a)$$

$$\Psi^+(x, y) = G'_{\sigma_y}(y - \epsilon) G_{\sigma_x}(x). \quad (20b)$$

This operator is specific to positive contrast lines with local orientation $\theta = 0$.

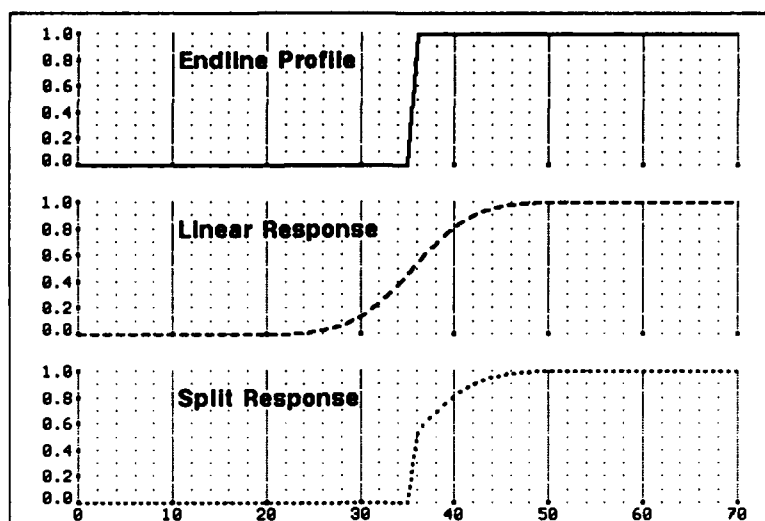


Figure 6: The signal (top) is the parallel cross-section of an image line near the discontinuous termination of the line (the endline). Note that the linear operator (middle) exhibits a smooth attenuation of response around the line ending. The goal (bottom) is an operator whose response attenuates abruptly at or near the endline discontinuity.

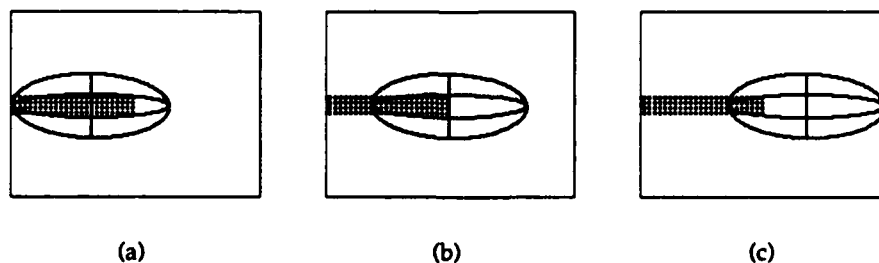


Figure 7: Schematic of the half-field decomposition and line endings. The elliptic regions in each figure represent the operator positions as the operator moves beyond the end of an image line. In (a) the operator is centered on the image line and the line exists in both half-fields. In (b) the operator is centered on the end-point and whereas the line only exists in one half-field, the other half-field contains the end-point. In (c) the operator is centered off the line and the line only exists in one half-field.

Note that this operator assumes both local continuity of α , and slow variation in contrast along the tangential direction. The indiscriminate smoothing of the gaussian G_{σ_x} provides a *post hoc* guarantee of these assumptions by smoothing out variation in the image. It will remove both orientation and contrast discontinuities (see Figure 6) in the image curve, thus violating the third criterion of §1.1. The local continuity of the curve must be verified *prior* to smoothing.

To resolve this, consider the continuity of the curve $\alpha(s)$.

DEFINITION 10: The curve α is C^1 Continuous at s_0 iff

$$\lim_{s \rightarrow s_0^-} \alpha(s) = \lim_{s \rightarrow s_0^+} \alpha(s) = \alpha(s_0) \quad (21a)$$

$$\lim_{s \rightarrow s_0^-} \theta(s) = \lim_{s \rightarrow s_0^+} \theta(s) = \theta(s_0) \quad (21b)$$

Thus, if the extensions of a curve on either side of a point $\alpha(s_0)$ both exist and are similarly oriented, then the hypothesis of local C_1 continuity at s_0 is confirmed.

DEFINITION 11: An operator is End-Line Stable if it responds positively only when centered on a continuous, connected region of an image curve.

Representing the intensity variation along the curve α as a function of the arc-length $I_\alpha(s)$, an ideal line-ending (or beginning) is a step in intensity at $s = 0$. End-line stability requires that the operator's response $\Psi(I_\alpha)|_s$ is non-positive for all $s \leq \epsilon$, and positive for $s > \epsilon$.^{*} Figure 7 demonstrates that this can be achieved by separately considering the behaviour of the curve in both directions around the operator centre.

We therefore adopt a partition which divides the operator kernel into regions along its length. Using the step function

$$\Sigma(x) = \begin{cases} 1, & \text{if } x > 0; \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

a partition of $F(x)$ around 0 is given by

$$F^-(x) = F(x) \Sigma(-x) \quad F^+(x) = F(x) \Sigma(x) \quad (23)$$

Note that $F^-(x) + F^+(x) = F(x)$ for all x , as required. A smooth (and more stable) partitioning operator is described in Appendix A.

Applying this partition to the gaussian smoothing kernel, we obtain the two half-field operators (see Figure 8(a))

$$G_\sigma^-(x) = G_\sigma(x) \Sigma(-x) \quad G_\sigma^+(x) = G_\sigma(x) \Sigma(x) \quad (24)$$

Which are combined to produce the end-line stable \hat{G} operator.

THEOREM 5: The operator $\hat{G}_\sigma = G_\sigma^- \Delta G_\sigma^+$ is end-line stable.

^{*}This property must also operate symmetrically at the other end of the curve.

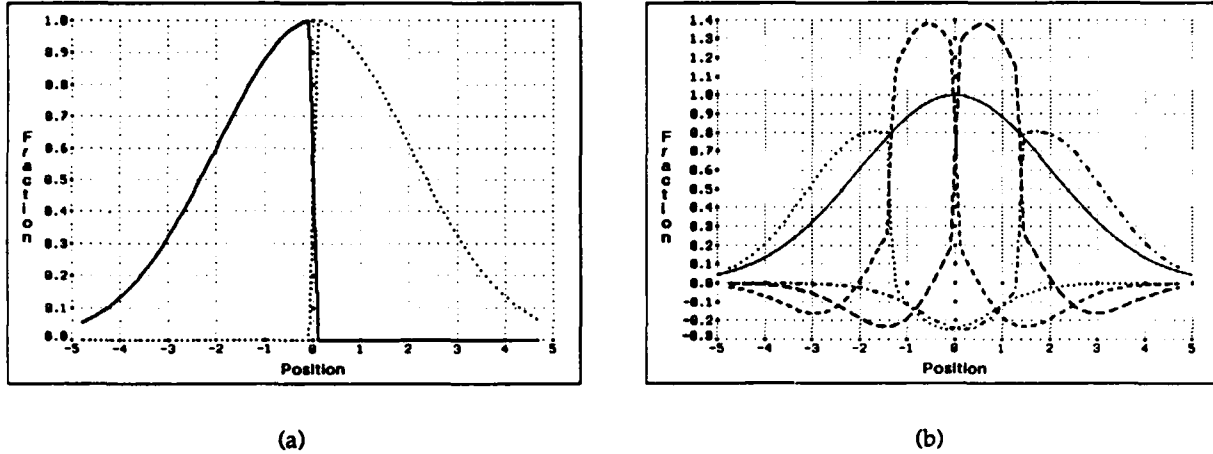


Figure 8: Splitting of 1D gaussian operator into components. Figure (a) shows the division about the center described by (24) and the subsequent analysis. Figure (b) is an example of the extension of this to four regions and the addition of response stabilizers as described in (27). The centers of division (c_i) partition the underlying area equally. The stabilizers have $\sigma' = 1.2$ and $w = 1/2$.

Consider the component responses in the neighbourhood of the step edge $I_\alpha(s) = \Sigma(s)$. The response of G_σ^+ to this step is given by

$$\begin{aligned}
 G_\sigma^+(I_\alpha)|_s &= (G_\sigma^+ * \Sigma)|_s \\
 &= \int_{-\infty}^{\infty} G_\sigma^+(s - \tau) \Sigma(\tau) d\tau \\
 &= \int_0^{\infty} \Sigma(s - \tau) G_\sigma(s - \tau) d\tau \\
 &= \int_{-\infty}^s \Sigma(\tau) G_\sigma(\tau) d\tau \\
 &= \begin{cases} \int_0^s G_\sigma(\tau) d\tau, & \text{if } s > 0. \\ 0, & \text{if } s \leq 0; \end{cases} \quad (25)
 \end{aligned}$$

The L/L AND of G_σ^+ and G_σ^- to produce \hat{G}_σ requires that both component responses be strictly positive for a positive response, thus whenever $s \leq 0$ around the step described above, the \hat{G} response is also zero. It is obvious that the same analysis applies to the symmetric G_σ^- component and the $1 - \Sigma(s)$ step edge, which describes behaviour around the other end of the line. Thus the \hat{G} operator is end-line stable symmetrically around a step edge. ■

Observe however that the component responses of the pure partition described above are unstable with respect to the Δ combinator. The $s \leq 0$ case in equation (25) produces a zero response, which forces the aggregate response of the Δ to be zero. The proximity of this value to the singular point of the Δ combinator ensures that the overall response is very sensitive to perturbations in the component responses. This can be ameliorated by adding a *stabilizer* to the partition, which will force the component responses negative where they would otherwise be simply zero. While many stabilizers will do we adopt the first derivative of a gaussian as perhaps

the "simplest" smooth operator whose response is entirely negative on one side of a step edge and entirely positive on the other.

This result can be combined with the approach of [Davis *et al.*, 1976], in which a linear operator is made noise-insensitive by dividing it into a number of sections along its length and 'detecting' a line only when more than half of the regions give a positive response (this is also similar to the ANDing of LGN afferents proposed in [Marr and Hildreth, 1980]). Combining these two principles requires that the operator be smoothly divided into some even number of regions around its centre. The partition of (23) can be generalized to $n > 2$ regions by combining partitions for a strictly increasing sequence of partition points $\{c_i \mid 1 \leq i < n\}$ ($S_i(x)$ is the stabilizer described above).

$$S_i(x) = w G'_\sigma(x - c_i) \quad (26)$$

$$F_i(x) = \begin{cases} F(x) (1 - \Sigma(x - c_1)) - S_i(x) & \text{if } i = 1; \\ F(x) (\Sigma(x - c_{i-1}) - \Sigma(x - c_i)) + S_{i-1}(x) - S_i(x) & \text{if } 1 < i < n; \\ F(x) \Sigma(x - c_{n-1}) + S_{n-1}(x) & \text{if } i = n. \end{cases} \quad (27)$$

Algebraic combination will verify that these F_i partition the function $F(x)$. The separation of a gaussian into four equal area components with G' stabilizers (which we will refer to as the *length components* of the operator) is shown in Figure 8(b).

Consider a gaussian operator thus partitioned into an even n components G'_σ by the introduction of $n - 1$ partition points c_i .^{*} Require then that all responses in one half-field and any response in the other half-field be positive.[†] This condition can be stated as

DEFINITION 12: (LENGTH COMBINATION 1) *To ensure that the operator \hat{G}_σ^* will only respond positively when centered on a locally continuous image curve, combine the length components of the operator as follows*

$$\hat{G}_\sigma^* = ((G_\sigma^1 \Delta \dots \Delta G_\sigma^{n/2}) \Delta (G_\sigma^{n/2+1} \nabla \dots \nabla G_\sigma^n)) \nabla ((G_\sigma^1 \nabla \dots \nabla G_\sigma^{n/2}) \Delta (G_\sigma^{n/2+1} \Delta \dots \Delta G_\sigma^n)). \quad (28)$$

Note that this combination of the length components is end-line stable (it is just a generalization of the 2-component case of Theorem 5), and that the noise insensitivity argument of [Davis *et al.*, 1976] still holds, since the conditions of (28) require that *more than half* of the G'_σ regions exhibit a positive response.

Although (28) is useful for verifying end-line stability, a logically more restrictive and simpler form is used for actually computing the operator responses.

DEFINITION 13: (LENGTH COMBINATION 2)

$$\hat{G}_\sigma = G_\sigma^{n/2} \Delta G_\sigma^{n/2+1} \Delta ((G_\sigma^1 \Delta \dots \Delta G_\sigma^{n/2-1}) \nabla (G_\sigma^{n/2+2} \Delta \dots \Delta G_\sigma^n)) \quad (29)$$

^{*}By spacing the partition points such that the area under each of the components G'_σ is equal, the contribution of each component will have the same significance on the final result. For a gaussian operator the partition points to produce four equal area components should be $\{-0.674\sigma, 0, 0.674\sigma\}$ and for six $\{-0.967\sigma, -0.431\sigma, 0, 0.431\sigma, 0.967\sigma\}$.

[†]There are a number of choices of logical combination of these responses which guarantee end-line stability, including the [Davis *et al.*, 1976] condition itself. This one is perhaps the most structurally restrictive while retaining logical simplicity.

This formulation simplifies the computation since each component appears exactly once.*

In order to complete the analysis, we unify this tangential continuity combination with the cross-sectional combination. Each of the responses $G_o^i(I)$ is an L/L combination of the i th length components of (20). Thus each length component is a test of the lateral preconditions, and the tangential combination tests for continuity ensuring end-line stability.

DEFINITION 14: The Logical/Linear Operators Ψ are given by

$$\Psi = \Psi_{n/2} \Delta \Psi_{n/2+1} \Delta \left((\Psi_1 \Delta \cdots \Delta \Psi_{n/2-1}) \nabla (\Psi_{n/2+2} \Delta \cdots \Delta \Psi_n) \right) \quad (30)$$

where

$$\begin{aligned} \Psi_i^- &= \Psi_i^- \Delta \Psi_i^+ && \text{for Positive Contrast Lines,} \\ \Psi_i &= -\Psi_i^- \Delta -\Psi_i^+ && \text{for Negative Contrast Lines,} \\ \Psi_i &= \Psi_i^* \Delta (\Psi_i^* - \Psi_i^-) \Delta (\Psi_i^* - \Psi_i^+) && \text{for Edges.} \end{aligned}$$

The length components Ψ_i^- , Ψ_i^+ are obtained by substituting (20) into (27). The Ψ^* and Ψ_i^* are obtained by a similar extension of (18a) to two dimensions and length components.

3.3 Curvature and End-Stopping

Extending the L/L operators to show curvature as well as orientation selectivity depends on the work of [Dobbins *et al* 1987; 1989], which describes a computational model of end-stopping in the primary visual cortex. The key finding is that end-stopping is related to the measurement of local curvature, and that models of end-stopped neurons can be derived by taking differences between oriented linear operators. This model can easily be expressed in the language of L/L operators, and can incorporate the stable, L/L operators developed above as components.

Given two similarly oriented line operators S (short) and L (long) and weighting factors to compensate for the differences in their relative areas (w_S and w_L), the end-stopped operator ES is given by

$$ES = \phi(w_S \phi(S) - w_L \phi(L)) \quad (31)$$

where $\phi(x)$ is a half-wave rectification. If the excitatory and inhibitory components are closely matched in both spatial frequency bandwidth (in the normal direction) and orientation bandwidth, then the ES operator will have a characteristic maximum response to a curved line with some non-zero curvature[†] (see Figure 9(a)). Since we wish to develop operators with signed curvature specificity, we will use an L/L line operator as the excitatory component and an L/L length-partitioned operator with a G' cross-section as the inhibitory component. This component is laterally offset from center so that its zero axis coincides with that of one lateral component of the excitatory operator; this places the positive part of the inhibitory component directly under the center of the excitatory operator, thus guaranteeing maximal inhibition for straight lines. The odd-symmetry of the inhibition in such an operator ensures that it inhibits response for only one sign of curvature (see Figure 9(b)), and thus results in an ES line operator which is tuned for both magnitude and sign of curvature.

*It should be obvious that a condition for reconstruction of a partitioned linear operator is that each component in the decomposition must appear exactly n times in the reconstruction expression.

[†]In all discussion of curvature values in this paper, curvature is expressed in units of 1/pixels. A curvature of 0.1 is therefore associated with a circle of radius 10 pixels.

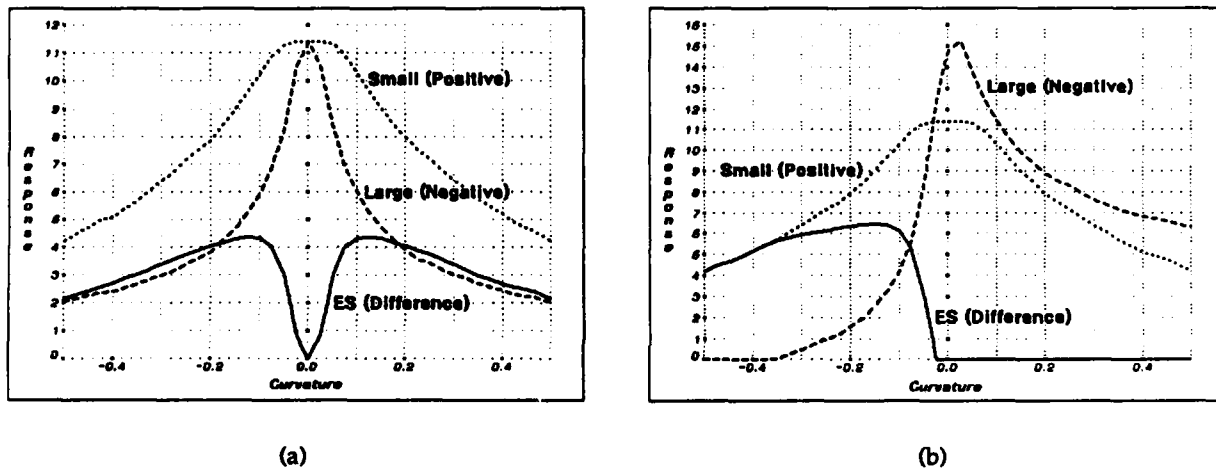


Figure 9: Curvature responses of operators resembling "simple cells" of two different lengths, and their difference, used to construct an end-stopped (ES) cell: (a) shows the responses for an even-symmetric inhibitory component, which leads to magnitude selectivity only; (b) shows that for an odd-symmetric inhibitory component, both sign and magnitude of curvature are selected.

4. Results

The parameters for three base operators are shown in Table 1. The non-endstopped operator and the excitatory part of the end-stopped operators are laterally decomposed with $a\sigma = \sigma_y$ and $b\sigma = \epsilon$ [see (20)]. In addition the tangential decomposition of these parts into n regions [see (27)] includes *stabilizers* with a $\sigma' = \sigma_x/4$, and $w = W$ [see (26)]. The inhibitory part of the end-stopped operators has an odd-symmetric G' cross-section laterally displaced by ϵ and is tangentially decomposed in the same way. The weights W are the weights on each part required to normalize the response for an ideal stimulus (%100 contrast line of width 2 pixels). The positive curvature operators are simply 180 degree rotations of the negative curvatures.

Comparing the operator responses to the test image in Figure 11 is revealing. All of the L/L operators are more specifically tuned than the linear (Fig. 12(a)), and each variation increases the operator selectivity. The cross-sectional L/L decomposition does indeed serve to localize responses laterally (Fig. 12(b)), although its strength in suppressing pure edge responses is not apparent here. The tangential L/L decomposition (Fig. 12(c)) does prevent the indiscriminate bridging of gaps between nearby curves, and the tendency to run off the ends of curves, two characteristics of the linear operator. Finally, the combination of these two decompositions produces an operator which behaves exactly as it should behave, with accurate localization in all directions of variation and stable behaviour in the vicinity of line endings and multiple lines.

Note that the only effect of the L/L decomposition is to eliminate certain of the responses from the linear operator. These examples show that the L/L have been effective in eliminating all but the 'best' from among the 'potential' responses. The standard response to this problem with linear operators (e.g. [Canny, 1986]) relies on arbitrary thresholding and further post-processing, and tends to leave the question of how to choose appropriate thresholds open to the user.

Compare the L/L edge measurements in Figure 18(a) with Canny's operator in Figure 18(b) (both processes are run at the same scale). Not only is the L/L edge operator as accurate, it makes explicit a great deal of information which is either unavailable or confounding in the Canny image. For example, the cross in the upper left is made by two straight lines crossing each other. This relationship is explicit in the multiple orientations exhibited at or near the crossing, whereas the Canny responses would indicate that there is, in fact, a smooth transition from vertical to horizontal in this neighbourhood. In addition, the contour following of the Canny process is not only a sequential bottleneck in what should be a purely parallel process, but also ensures that the 'noise' responses appear as well structured as the 'true' responses, thus interfering with the potential for a later process to effectively discriminate between them. And again we are faced with the same question as with the line example discussed above—what threshold should one

	ES?	ES Component	σ_x	σ_y	ϵ	W
$ \kappa = 0.0$	No		3.0	1.0	1.0	0.175
$ \kappa = 0.1$	Yes	Excitatory	2.0	1.0	1.0	0.242
		Inhibitory	3.0	1.25	1.25	0.316
$ \kappa = 0.2$	Yes	Excitatory	1.25	1.0	1.0	0.245
		Inhibitory	2.25	1.25	1.25	0.490

Table 1: Parameters for a family of L/L operators

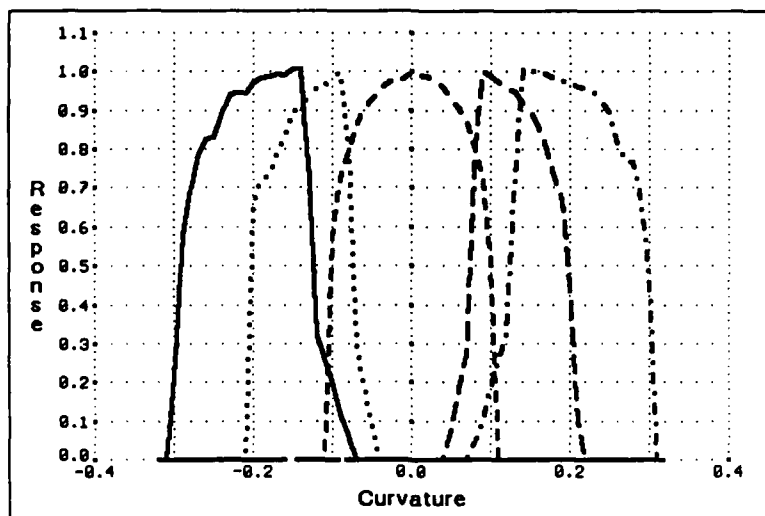


Figure 10: Response profiles of a family of curvature-selective line operators tuned for curvatures $\{-0.2, -0.1, 0.0, 0.1, 0.2\}$. These operators are described in Table 1.

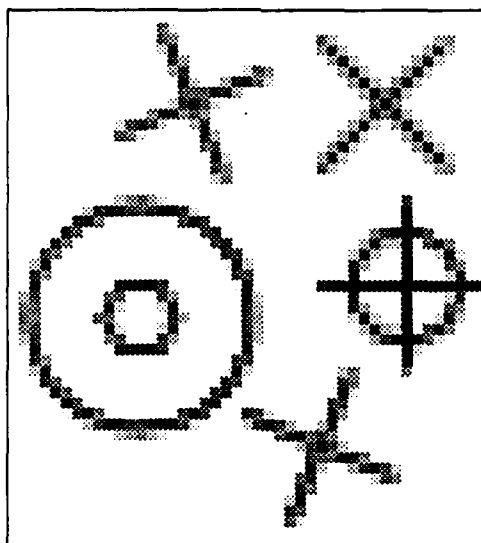


Figure 11: An artificial image used to test the image operators. This is an anti-aliased grey-scale image of lines and curves, which represent the range of orientations and curvatures represented by an operator family and the singularities (line endings and crossing lines) which generally cause problems for linear operators.

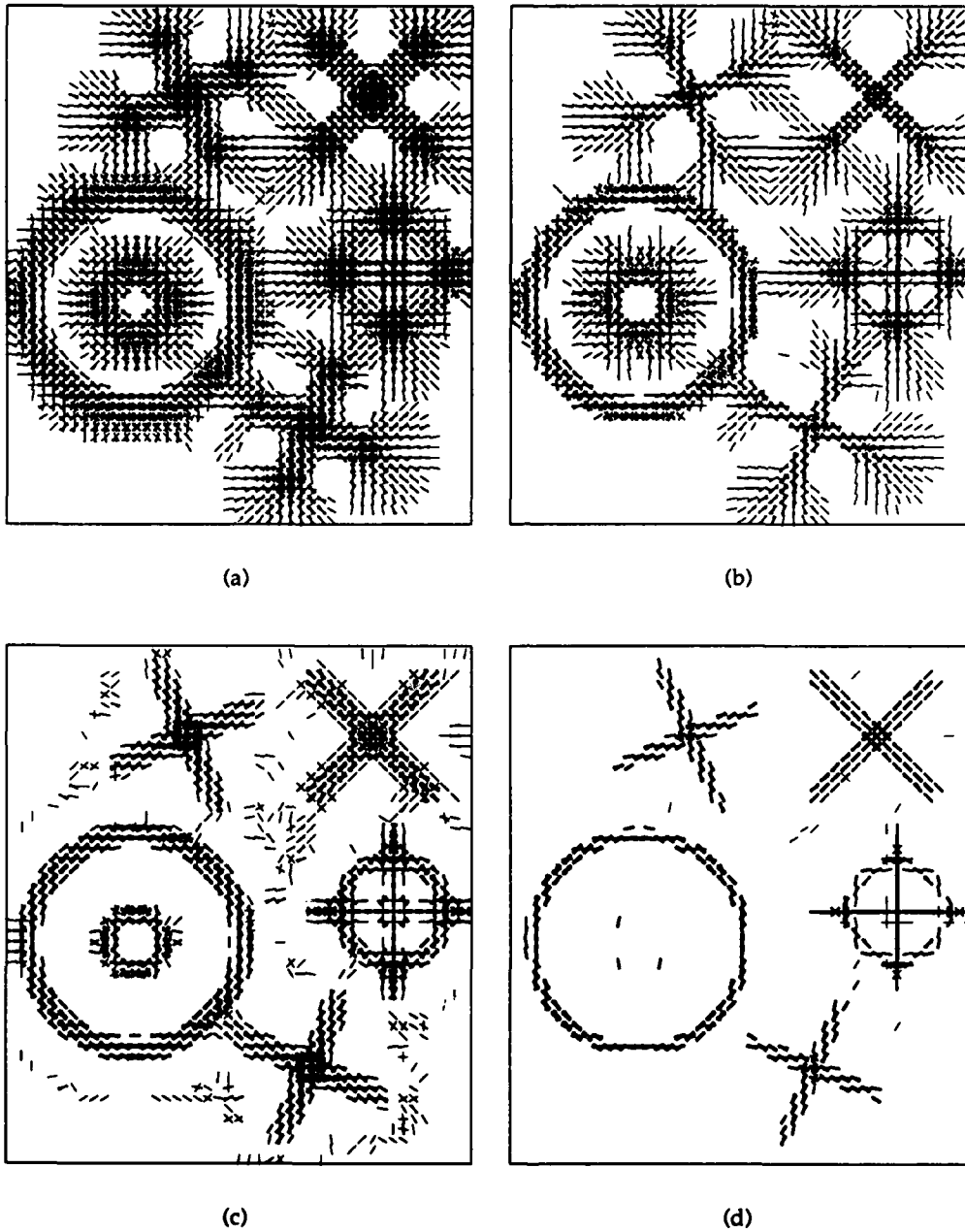


Figure 12: The operator responses for the artificial test image. For each test, the responses shown are for 12 orientations of a non-end-stopped simple cell. (a) the positive operator responses for a pure linear operator, (b) the responses for an operator with only the cross-sectional non-linearity, (c) the responses for an operator with only the tangential non-linearity, and (d) the responses for the operator with both non-linearities.

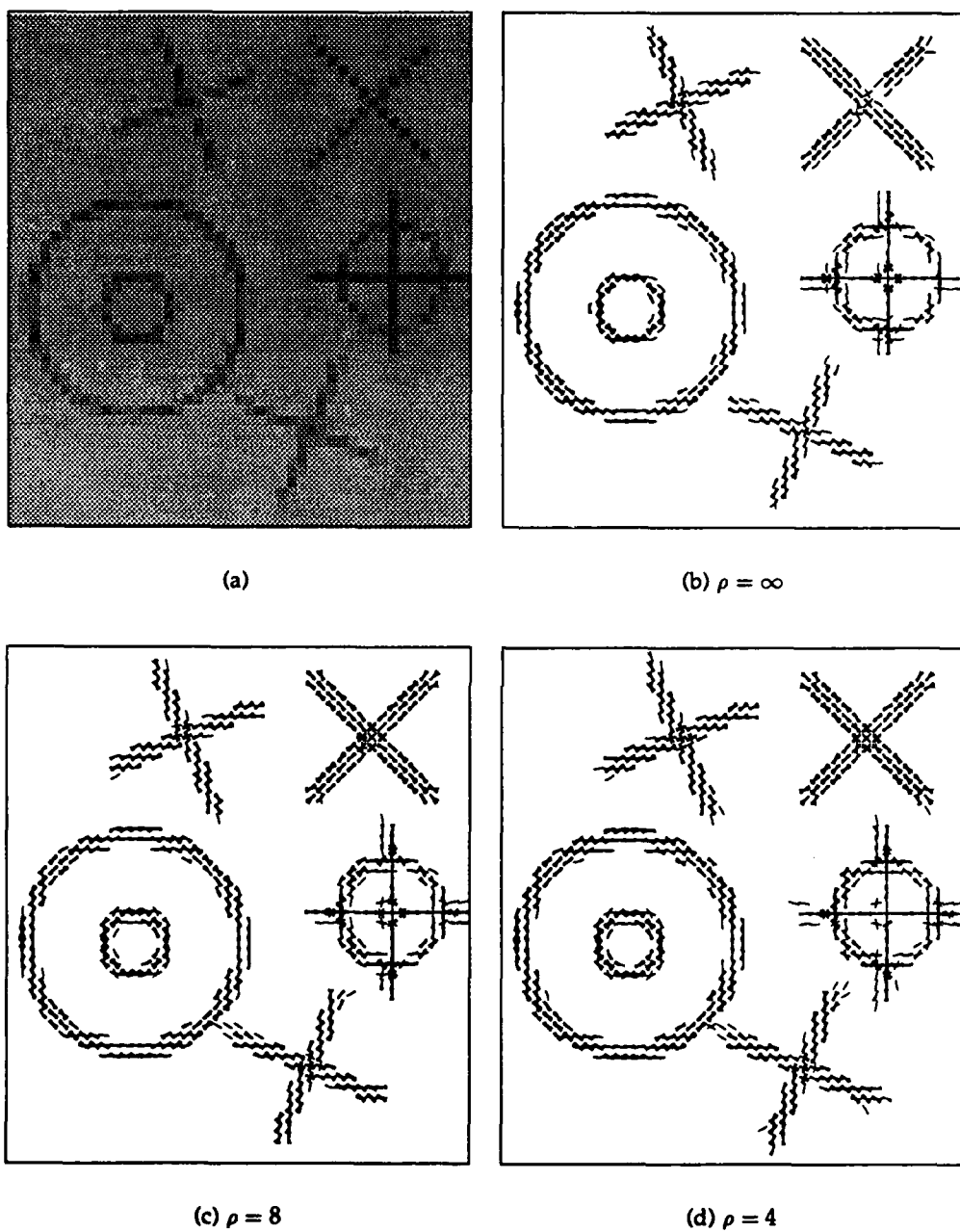


Figure 13: The effect of varying the L/L ρ on a low-contrast version of the artificial test image. (a) Shows the low-contrast image, and (b) the operator responses with ideal L/L operators $\rho = \infty$, (c) responses for $\rho = 8$, and (d) for $\rho = 4$.

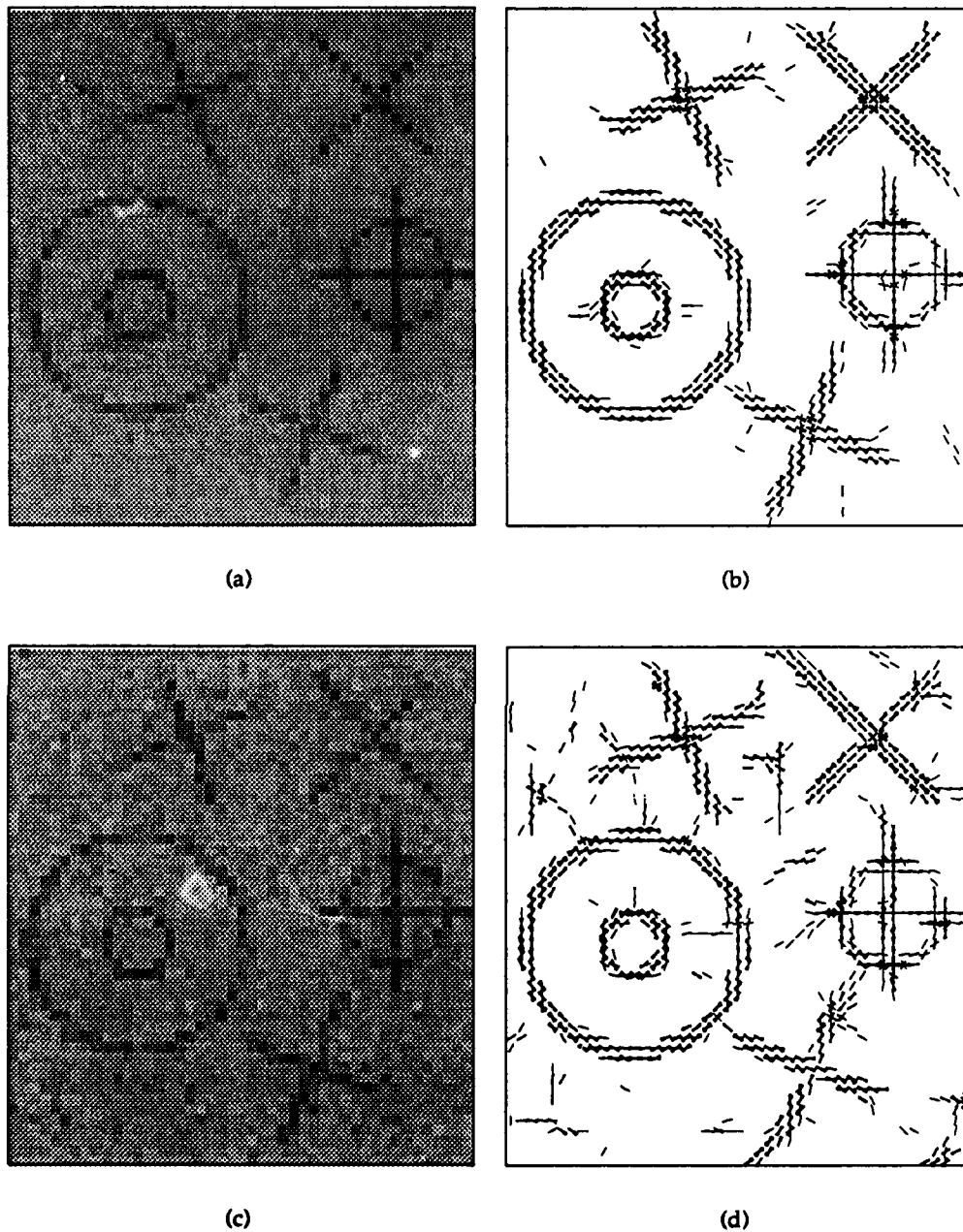


Figure 14: The effect of adding gaussian noise to a low-contrast version of the artificial test image. (a) Shows the image with additive noise of $\sigma = 1.0$, and (b) the operator responses. (c) Shows the image with additive noise of $\sigma = 2.0$, and (d) the operator responses.

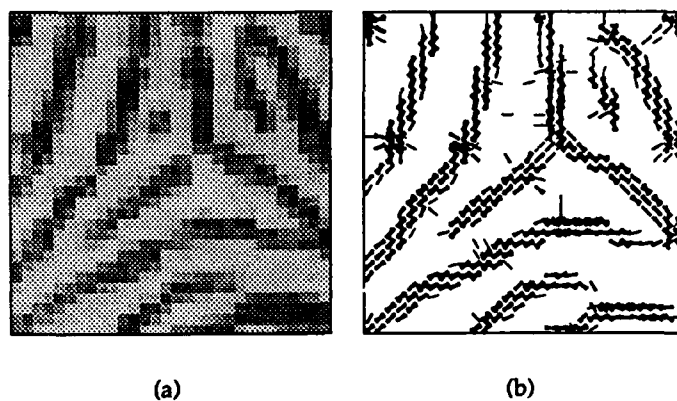


Figure 15: L/L responses to a fingerprint image.

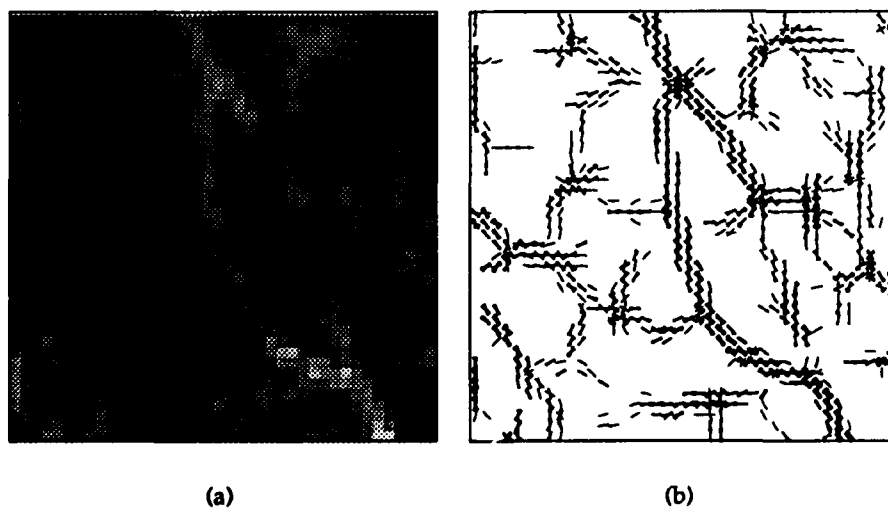
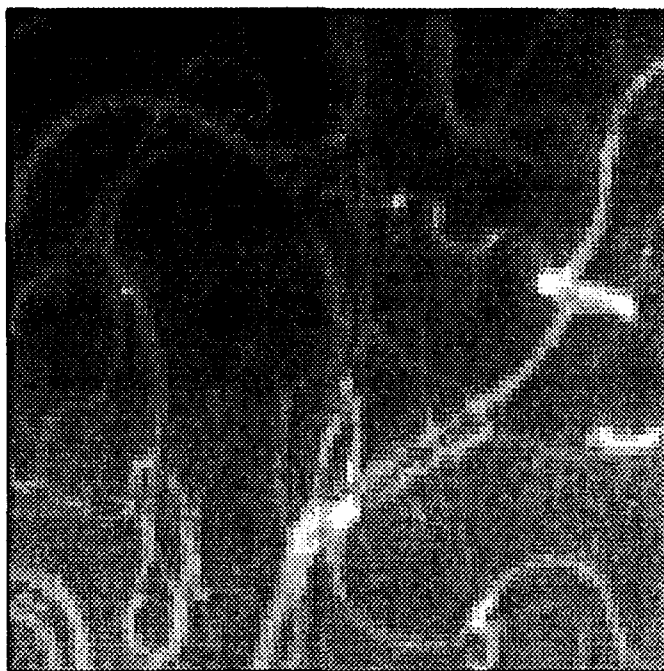
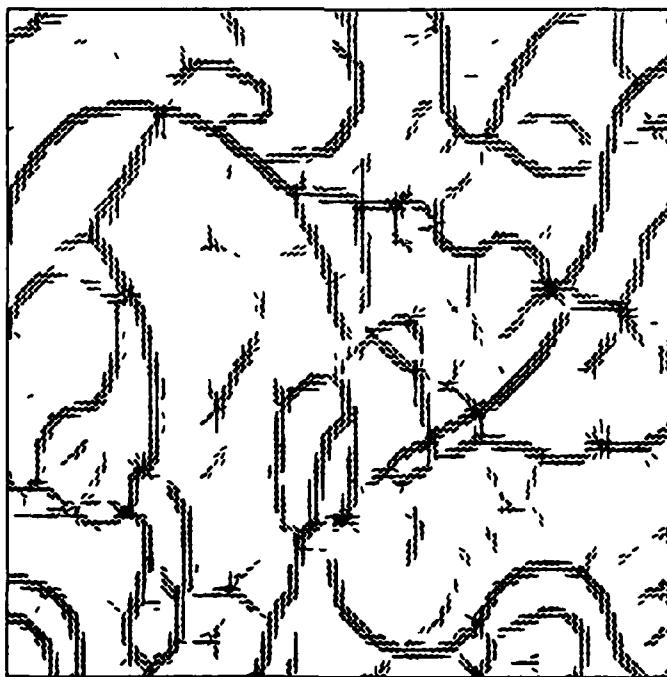


Figure 16: L/L responses to a remote sensing image of logging roads.

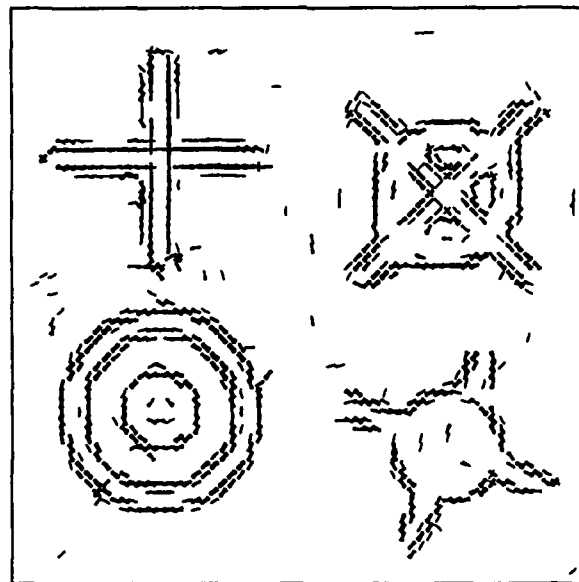


(a)

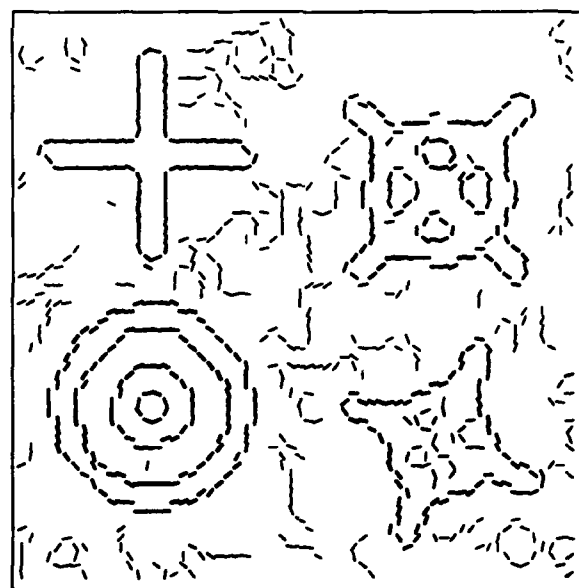


(b)

Figure 17: L/L responses to a cerebral angiogram.



(a)



(b)

Figure 18: Comparison of edge responses for a noisy test figure and (a) the L/L edge operators, (b) Canny's operator.

use? Of course, the Canny responses could be 'cleaned-up' by making the high threshold larger, but at what price and on what criteria?

The obvious strengths of the examples shown above are an accurate estimate of the positions, orientations and curvatures of the image curves. Discontinuities and bifurcations in these curves are not smoothed over, instead they are represented as multiple tangents at the same point. This is clearest in the natural images. However, it is clear that these measurements are still somewhat sensitive to noise and local perturbations, and do not always provide the best estimates of local curvature.

5. Discussion

5.1 Biological Implications

A number of observations regarding the relationships between this model and natural visual systems are possible. In general, there is nothing in the operator structure which is biologically *implausible*; in fact, some attempt has been made (see Appendix A) to demonstrate how such a model might be instantiated in neurophysiological terms. The gross similarities between these oriented, contrast-tuned operators and cortical simple cells have already been noted. The model of end-stopping which has been adopted to give these operators curvature-sensitivity was motivated largely as a biological model. However, a number of more intriguing observations can be made.

Evidence was obtained that the tangential non-linearities proposed may have a neurophysiological correlate by [Hammond and MacKay, 1983]. In doing single-cell recordings using illuminated bars as stimuli, they observed that a small opposite contrast region in the midst of a stimulus was able to inhibit the response of a simple cell much more effectively than would be predicted by a linear model. The opposite contrast region was, in effect, able to turn the cell off. This is exactly the behaviour caused by the tangential partitioning of the operator.

More recently, [Bolz *et al.*, 1989] observed facilitatory interactions between similarly oriented layer 5 and layer 6 pyramidal cells. The layer 5 cells had small receptive fields (RF's) which all lay within the RF of the layer 6 cell. By inhibiting the activity of one or more of the layer 5 cells, the length response of the layer 6 cell was reduced by the size of the inhibited RF. They concluded that the layer 6 cells were "constructed" by an additive summation of the shorter layer 5 cells. This description corresponds exactly to the model described in this paper.

Potentially problematic is the work of [Schumer and Movshon, 1984], which discredits the [Marr and Hildreth, 1980] AND-gating model, and can also be extended to the current analysis. Schumer and Movshon demonstrate that there is no real minimum length of bar stimulus required to stimulate a simple cell, instead the minimum length required to produce a noticeable stimulation is inversely related to stimulus contrast. While this is certainly a disproof of the very simplistic thresholded AND-gating model of Marr and Hildreth, it is not inconsistent with the model we are proposing. To see this, one must pay close attention to the consequences of the smooth approximations to the L/L AND and OR operators described in Appendix A. For any value of ρ less than ∞ (the "hard" Δ), the consequence of the combination of a *zero* input with a positive input is to reduce but not eliminate the output ($|x| \Delta_{\rho} 0 = |x|/2$). Thus if all tangential regions but one are giving zero responses, the effect will be to significantly reduce the response of the L/L operator. The significant observation is that the "soft" L/L combinators within the RF summation network are sensitive to changes in *contrast-polarity* but insensitive to changes in *contrast-magnitude*. The analysis of [Schumer and Movshon, 1984] serves only to discredit the use of contrast-magnitude sensitive logical length summation models, their observations are entirely consistent with the L/L model.

The reader may have noticed that in Figure 5(c) the response region is offset from the actual peak of the 1D signal by a small amount. In light of the insistence that the operator respond only at peaks, this offset may seem problematic until the significance of an observation made in [Watt and Morgan, 1983; Whitaker and Walker, 1988] is recognized. It was observed that when a discrimination task involved locating the position of dot clusters or lines with non-constant intensity cross-section, the locations chosen were best described as the centroids of the intensity

distributions. The behaviour seen with our model is consistent with this interpretation, as the centroid of the triangular region is ≈ 56 which is definitely inside the region of positive response. To see why the model should exhibit this behaviour, notice that the peak of a gaussian convolution with a positive signal approaches the centroid of the signal as the width of the gaussian increases. As long as the gaussian is wide with respect to the signal, the peak of its response (as determined by the peak detection property of the line operator) will be close to the centroid of the intensity distribution. This relationship should be investigated further.

A number of more specific comparisons, relating our model to neurophysiological and psychophysical observations were explored in [Dobbins *et al.*, 1990], and will be expanded in future work.

5.2 Neural Networks

Any system whose behaviour depends on a linear summation of input values can be redesigned to take advantage of the properties of the L/L combinators. Candidates other than the vision operators already discussed include many "neural network" algorithms.

Consider, as a typical example, the feed-forward networks described in [Rumelhart *et al.*, 1986a]. Here an *activation function* F is applied to a weighted sum of inputs o to generate the updated activation a_i^{t+1} for each node i in the network. Thus

$$a_i^{t+1} = F\left(\left(\sum_j w_{ij} o_j^t\right) - \theta_j\right) \quad (32)$$

The weighted sum in this expression is just a one-dimensional discrete convolution over the previous output values o^t . Therefore, an L/L decomposition of this summing network can lead to performance improvements analogous to those in our image operators.

We have already applied L/L decomposition to a related *relaxation labelling* system [Hummel and Zucker, 1983; Iverson, 1988] for exposing patterns of consistency in our image operator responses (i.e. to suppress noise and impose a more specific model of curve continuity). In this particular system, the addition of L/L decompositions of the support network allows for successful and stable convergence of the relaxation in as few as *two or three iterations*. The details of this application will be discussed in a companion paper.

5.3 Conclusions

One of the major problems with linear operator approaches to detecting image curves is their false-positive responses to uncharacteristic stimuli. After outlining the necessary logical conditions for the existence of an image curve, we developed an operator decomposition which allows for the efficient testing of these conditions, and the elimination of these false-positive responses. To achieve this, it was essential to consider both the cross-section of the intensity image and the low-order differential structure of the curve itself.

The model appears consistent with both psychophysical and neurophysiological data and indicates non-linear interactions between simple cells in the inter-blob regions of layers 4, 5 and 6 of striate cortex. The model also suggests functional roles for dendro-dendritic interactions and shunting inhibitions in the dendritic trees of these cells.

More generally, we have introduced a flexible language for describing a useful class of non-linearities in operators. This language of *logical/linear* operators serves to combine existing linear models with logical descriptions of structure to produce operators which have guaranteed

stable behaviour. This class of operators represents a new approach to the problem of translating linear measures into logical categories. Thus they may prove essential in the eventual solution of a wide variety of classification problems, and in the principled and realistic modelling of neural networks.

A. Logical/Linear Combinators

As was described in §2.1 a set of conditionally linear logical combinators have been developed. The logical structure of these combinators is based on the principle that the sign of a signal (positive or negative) can be associated with a truth value (true and false, respectively). If some logical condition on the inputs is found to hold then the output will be a linear combination of the inputs.

The examples outlined are the L/L combinators analogous to logical *and* and *or*, and the rectifying operators $|x|$ and $\phi(x)$. The notation adopted reflects the combination of linear and logical principles in these operators:

- logical/linear AND of x and y is $x \Delta y$;
- logical/linear OR of x and y is $x \nabla y$.

The definitions and proofs in §2 clarify the relationship between these operators and both logical and linear principles. The symmetry and associativity of these operators are also easy to demonstrate.

A useful formulation of these operators is expressed in terms of the step function $\sigma(x)$:

$$\sigma(x) = \begin{cases} 1, & \text{if } x > 0; \\ 0, & \text{otherwise.} \end{cases} \quad (33)$$

This function is a choice operator pivoting around zero, and as such it can be used to directly define the L/L combinators above. Consider that the operators can be defined as:

$$x \Delta y = (x \text{ unless } \{x > 0 \wedge y \leq 0\}) + (y \text{ unless } \{y > 0 \wedge x \leq 0\}) \quad (34a)$$

$$x \nabla y = (x \text{ unless } \{x \leq 0 \wedge y > 0\}) + (y \text{ unless } \{y \leq 0 \wedge x > 0\}) \quad (34b)$$

$$\phi(x) = x \text{ when } \{x > 0\} \quad (34c)$$

$$|x| = (x \text{ when } \{x > 0\}) + (-x \text{ when } \{x \leq 0\}) \quad (34d)$$

Verify that these are equivalent to the algebraic expressions

$$x \Delta y = x(1 - \sigma(x)\sigma(-y)) + y(1 - \sigma(y)\sigma(-x)) \quad (35a)$$

$$x \nabla y = x(1 - \sigma(y)\sigma(-x)) + y(1 - \sigma(x)\sigma(-y)) \quad (35b)$$

$$|x| = x\sigma(x) - x\sigma(-x) \quad (35c)$$

$$\phi(x) = x\sigma(x) \quad (35d)$$

These are derived based on the principle of selecting those cases in which either x or y contributes linearly to the output.

In order to derive an analytic form for these operators we adopt a smooth approximation to $\sigma(x)$ *

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0; \\ e^{-1/x}, & \text{otherwise.} \end{cases} \quad (36)$$

*The 'logistic' sigmoid function of [Rumelhart *et al.*, 1986b] is another option for this, but the fact that the function chosen is only non-singular (i.e. $0 < \sigma_\rho(x) < 1$) on $x \in [-1/\rho, 1/\rho]$ means that the "hard" logic of §2.1 still applies for values outside of this region.

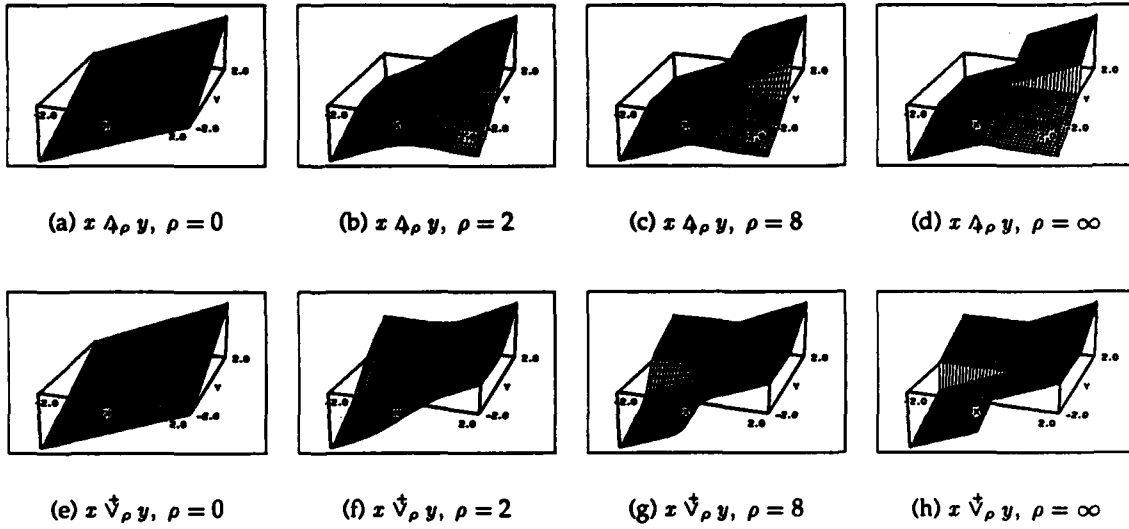


Figure 19: Graphs of L/L combinators varying ρ . (a), (b), (c), and (d) show $x \Delta_\rho y$, (e), (f), (g), and (h) show $x \nabla_\rho y$.

$$\sigma_\rho(x) = \frac{f(1/2 + \rho x)}{f(1/2 + \rho x) + f(1/2 - \rho x)} \quad (37)$$

This function is differentiable and

$$\sigma(x) = \lim_{\rho \rightarrow \infty} \sigma_\rho(x). \quad (38)$$

Furthermore this approximation makes clear the relationship between the linear sum and the L/L operators since $\sigma_0(x) = 1/2$. In this degenerate case, both of the operators simplify to a linear combination

$$\begin{aligned} x \Delta_0 y &= x(1 - \sigma_0(x) \sigma_0(-y)) + y(1 - \sigma_0(y) \sigma_0(-x)) \\ &= 3/4 (x + y) \end{aligned} \quad (39a)$$

$$\begin{aligned} x \nabla_0 y &= x(1 - \sigma_0(y) \sigma_0(-x)) + y(1 - \sigma_0(x) \sigma_0(-y)) \\ &= 3/4 (x + y) \end{aligned} \quad (39b)$$

Thus, the approximating functions Δ_ρ and ∇_ρ form a continuous deformation from a linear combination to the absolute L/L operation as ρ goes from 0 to ∞ .

A local network model of a Δ circuit is shown in Figure 20. This model is based on a simple principle, the *shunt* switch. Such a switch acts as a resistive element for its input signal with the resistance controlled by the control input. When the control input is negative, then the shunt is *on* and has a very low resistance; it passes its input through effectively unchanged. When the control is positive the shunt is *off* and the resistance is very high; the output is effectively zero. For simplicity, the shunts pictured allow multiple control inputs, all of which must be positive

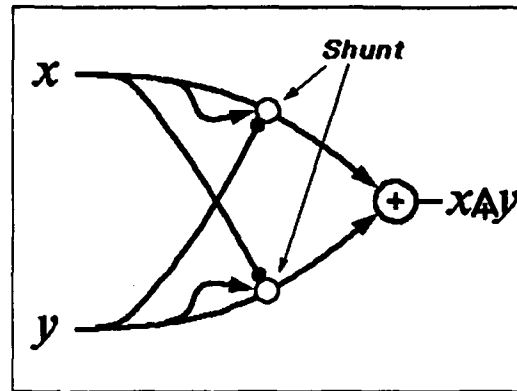


Figure 20: Network implementation of \mathbb{A} using a switching element. The shunts (small circles) operate by passing their input signal unchanged unless all of the control inputs are positive, in which case the output of the shunt is zero.*

for the shunt to turn off. This mechanism is similar to the functional characteristics of a relay or transistor. In neurophysiological terms, it could be related to shunting inhibition [Fatt and Katz, 1953; Rall, 1964] in the dendritic tree of a single neuron. The controls in this case would be realized as axo-dendritic or dendro-dendritic synapses.

Bibliography

- [Birkhoff and MacLane, 1977] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra (fourth edition)*. New York, N. Y.: Macmillan Publishing Co., Inc., 1977.
- [Bolz *et al.*, 1989] J. Bolz, C. D. Gilbert, and T. N. Wiesel, "Pharmacological analysis of cortical circuitry," *Trends in Neurosciences*, vol. 12, no. 8, pp. 292–296, 1989.
- [Canny, 1986] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 679–698, 1986.
- [Davis *et al.*, 1976] L. Davis, A. Rosenfeld, and A. Agrawala, "On models for line detection," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, no. 2, pp. 127–133, 1976.
- [Deriche, 1987] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *International Journal of Computer Vision*, vol. 1, no. 2, pp. 167–187, 1987.
- [Dobbins *et al.*, 1987] A. Dobbins, S. W. Zucker, and M. Cynader, "Endstopping in the visual cortex as a substrate for calculating curvature," *Nature*, vol. 329, pp. 438–441, 1987.
- [Dobbins *et al.*, 1989] A. Dobbins, S. W. Zucker, and M. Cynader, "Endstopping and curvature," *Vision Research*, vol. 29, pp. 1371–1387, 1989.
- [Dobbins *et al.*, 1990] A. Dobbins, L. Iverson, and S. W. Zucker, "A logical/linear model of cortical subunit interactions," in *ARVO Annual Meeting: Abstract Issue*, (Sarasota, Florida), p. 397, May 1990.
- [Fatt and Katz, 1953] P. Fatt and B. Katz, "The effect of inhibitory nerve impulses on a crustacean muscle fibre," *J. Physiology (London)*, vol. 121, pp. 374–389, 1953.
- [Hammond and MacKay, 1983] P. Hammond and D. M. MacKay, "Influence of luminance gradient reversal on simple cells in feline striate cortex," *J. Physiology (London)*, vol. 337, pp. 69–87, 1983.
- [Herskovitz and Binford, 1970] A. Herskovitz and T. O. Binford, "On boundary detection," MIT AI Memo 183, MIT AI Lab, Cambridge, Mass., 1970.
- [Heuckel, 1971] M. H. Heuckel, "An operator which locates edges in digital pictures," *J. Association for Computing Machinery*, vol. 18, pp. 113–125, 1971.
- [Horn, 1977] B. K. P. Horn, "Understanding image intensities," *Artificial Intelligence*, vol. 8, pp. 201–231, 1977.
- [Hummel and Zucker, 1983] R. A. Hummel and S. W. Zucker, "On the foundations of relaxation labeling processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 267–287, 1983.
- [Iverson, 1988] L. A. Iverson, "Discrete forms of continuity in space: The description of image curves," Master's thesis, McGill University, 1988.

- [Jones and Palmer, 1987] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophys.*, vol. 58(b), pp. 1233-1258, 1987.
- [Koenderink and van Doorn, 1976] J. J. Koenderink and A. J. van Doorn, "The singularities of the visual mapping," *Biological Cybernetics*, vol. 24, pp. 51-59, 1976.
- [Koenderink and van Doorn, 1982] J. J. Koenderink and A. J. van Doorn, "The shape of smooth objects and the way contours end," *Perception*, vol. 11, pp. 129-137, 1982.
- [Koenderink and van Doorn, 1987] J. J. Koenderink and A. J. van Doorn, "Representation of local geometry in the visual system," *Biological Cybernetics*, vol. 55, pp. 367-355, 1987.
- [Koenderink, 1984] J. J. Koenderink, "The internal representation of solid shape and visual exploration," in *Sensory Experience, Adaptation, and Perception*, ch. 7, pp. 123-142, Lawrence Erlbaum Associates, Inc., 1984.
- [Leclerc and Zucker, 1984] Y. Leclerc and S. W. Zucker, "The local structure of image discontinuities in one dimension," Computer Vision and Robotics Lab Technical Report 83-19R, McGill University, Montreal, 1984.
- [Marr and Hildreth, 1980] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Royal Society*, vol. 207, pp. 187-217, 1980.
- [Marr, 1982] D. Marr, *Vision*. San Francisco: W.H. Freeman, 1982.
- [Movshon *et al.*, 1978] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst, "Spatial summation in the receptive fields of simple cells in the cat's striate cortex," *J. Physiology (London)*, vol. 283, pp. 53-77, 1978.
- [Protter and Weinberger, 1984] M. H. Protter and H. F. Weinberger, *Maximum Principles in Differential Equations*. New York, N. Y.: Springer-Verlag, 1984.
- [Rall, 1964] W. Rall, "Theoretical significance of dendritic trees for neuronal input-output relations," in Reiss [1964], pp. 73-97.
- [Reiss, 1964] R. Reiss, ed., *Neuronal Theory and Modelling*. Stanford, Cal.: Stanford Univ. Press, 1964.
- [Rumelhart *et al.*, 1986a] D. E. Rumelhart, G. E. Hinton, and J. L. McClelland, "A general framework for parallel distributed processing," in Rumelhart *et al.* [1986c], ch. 2, pp. 45-76.
- [Rumelhart *et al.*, 1986b] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in Rumelhart *et al.* [1986c], ch. 8, pp. 318-362.
- [Rumelhart *et al.*, 1986c] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, Mass.: The MIT Press, 1986.
- [Schumer and Movshon, 1984] R. A. Schumer and J. A. Movshon, "Length summation in simple cells of cat striate cortex," *Vision Research*, vol. 24, no. 6, pp. 565-571, 1984.
- [Shapley and Lennie, 1985] R. Shapley and P. Lennie, "Spatial frequency analysis in the visual system," *Annual Review of Neuroscience*, vol. 8, pp. 547-583, 1985.

- [Sikorski, 1960] R. Sikorski, *Boolean Algebras*. New York, N. Y.: Springer-Verlag, 1960.
- [Watt and Morgan, 1983] R. J. Watt and M. J. Morgan, "Mechanisms responsible for the assessment of visual location: theory and evidence," *Vision Research*, vol. 23, pp. 97--109, 1983.
- [Whitaker and Walker, 1988] D. Whitaker and H. Walker, "Centroid evaluation in the vernier alignment of random dot clusters," *Vision Research*, vol. 28, pp. 777--784, 1988.
- [Zucker *et al.*, 1988] S. W. Zucker, C. David, A. Dobbins, and L. Iverson, "The organization of curve detection: Coarse tangent fields and fine spline coverings," in *Second International Conference on Computer Vision*, (Tampa, Florida), pp. 568--577, IEEE Computer Society, December 1988.

Appendix II

A Logical/Linear Model of Cortical Subunit Interactions.

A LOGICAL/LINEAR MODEL OF CORTICAL SUBUNIT INTERACTIONS

Allan Dobbins Lee Iverson
Steven W. Zucker

Research Centre for Intelligent Machines,
McGill University, Montréal.

Abstract

Visual cortical neurons vary dramatically in their tuning along different stimulus dimensions, and in whether those selectivities are adequately captured by linear models. Certain properties, such as spatial frequency selectivity, seem adequately explained by linear models; while others, such as direction selectivity and endstopping, may involve nonlinearities. We have developed a calculus for the logical combination of local evidence from subunits in which support is accumulated linearly, but incompatible evidence enters nonlinearly. This leads to a class of operators that appears linear for one class of stimuli but markedly nonlinear for others — we call these logical-linear operators. They exhibit dual advantages: they are considerably more stimulus-specific than purely linear operators, while more robust to incidental stimulus variation than logical operators. The viability of such operators as models of visual cortical neurons (e.g. simple cells) is examined by comparing simulations of purely linear and logical-linear models to the responses of cortical neurons. Operators with properties specialised for spatial contours are examined with stimuli containing vernier offsets, interruptions, and opposite contrast segments. The results are consistent with the well-known “linear” properties (e.g., sensitivity to s.f. gratings) while exhibiting the nonlinear behaviour associated with high vernier sensitivities (Swindale and Cynader, 1989) and strong suppressive effects for opposite contrast segments (Hammond and MacKay, 1983, 1985). Finally, we relate the abstract description of subunit combination to nonlinear synaptic interactions.

1. Motivation

The goal of the visual system is to provide a description of the visual world, an essential feature of which is the accurate categorization of stimuli. Curves in particular require both existence/non-existence decisions and categorical distinctions (e.g. between lines and edges). Linear operators (modelling cortical cell RFs) are often assumed to be one of the first stages in such a system.

Linear operators are not decision systems however, and thus depend on post-processing (e.g. thresholding) to locate and describe curves. However, the indiscriminate combination of evidence characteristic of linear summation often destroys the very information necessary for accurate descriptions. Instead we propose an alternative approach whereby linear models are decomposed into structurally significant subunits which are thresholded only for *contrast polarity*. The resulting models display much higher structural selectivity than linear models, yet still appear linear in many standard experimental paradigms.

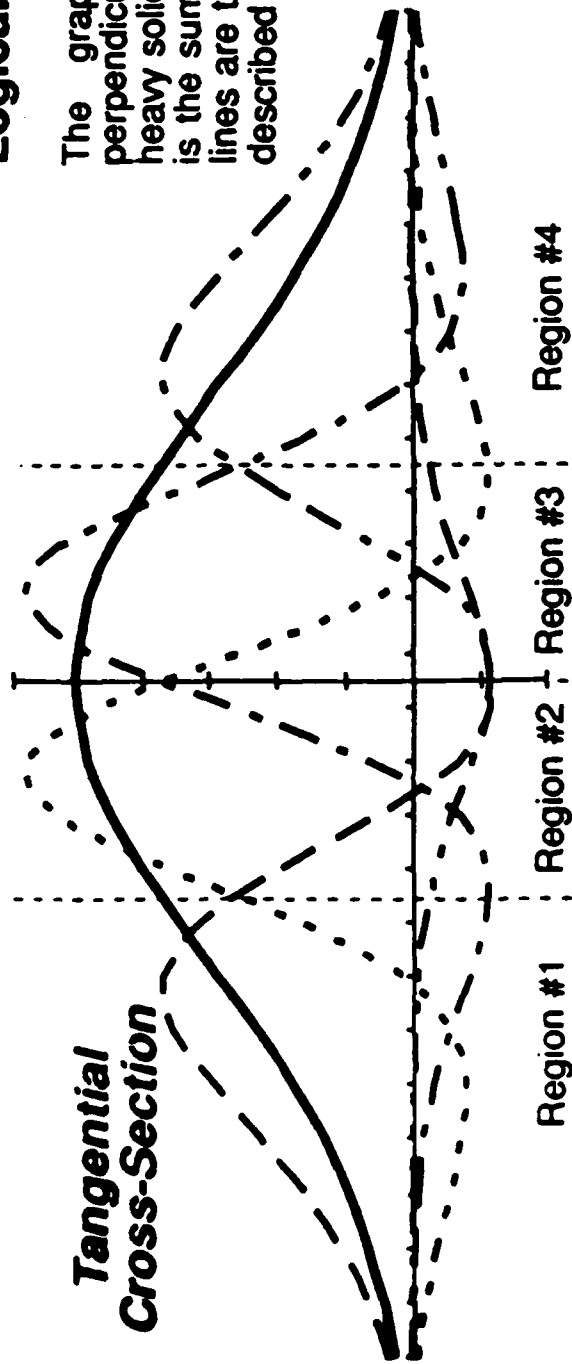
2. Subunit Decomposition

Linear simple cell RFs were modeled with a derivative of Gaussian model. In the simulations shown here the RF is an even symmetric second derivative of Gaussian.

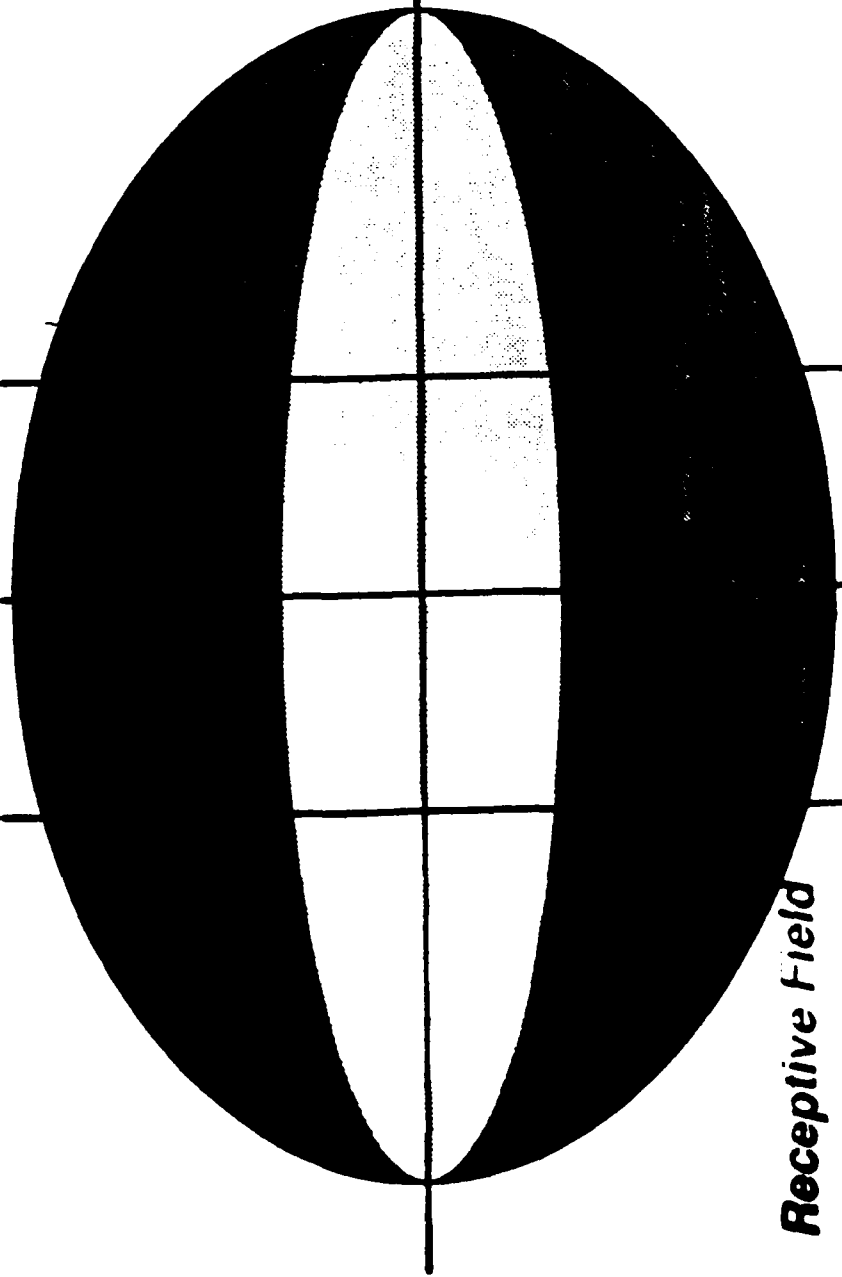
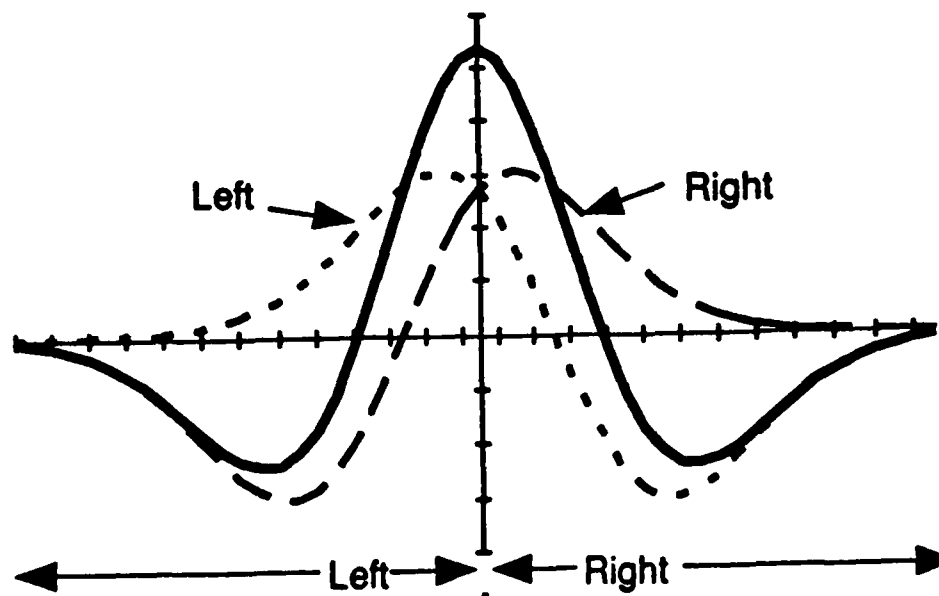
To derive the logical/linear models, the linear RF is decomposed into a set of subunits whose sum equals the original. Each subunit represents a component in a boolean decision tree specific to locally continuous positive contrast lines. The partition chosen for these simulations is mathematically convenient but the subunits do not correspond to observed RFs of cortical or geniculate cells in a straightforward way. However the behaviour depends principally on the geometry of the decomposition and the sign of local measurements and not on the detailed subunit form.

Logical/Linear Cell Schematic

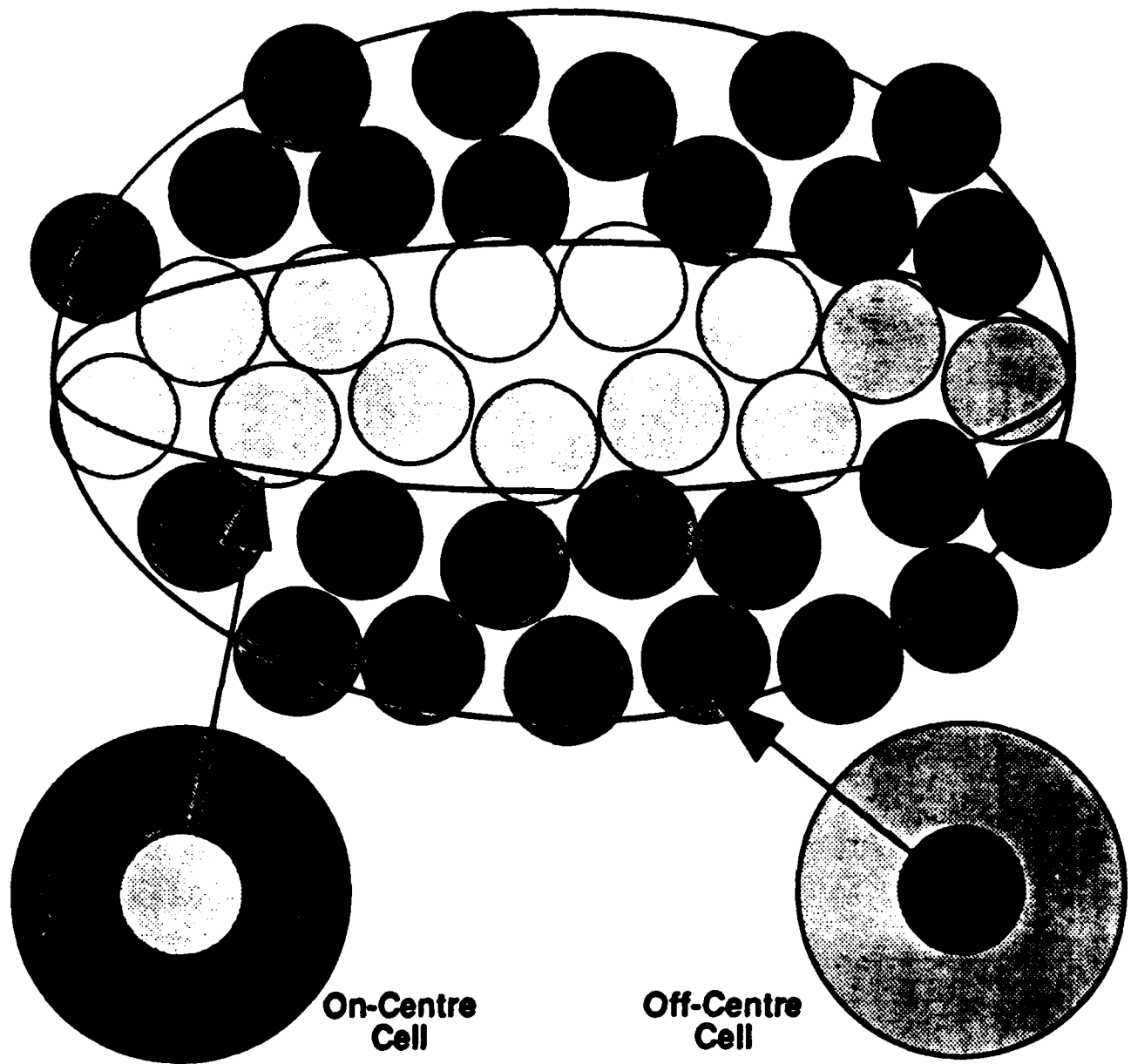
The graphs show the cross-section (in perpendicular directions) of the model. The heavy solid line represents the linear cell, which is the sum of the cross-sections. The dashed lines are the components of the decomposition described above.



Lateral Cross-Section



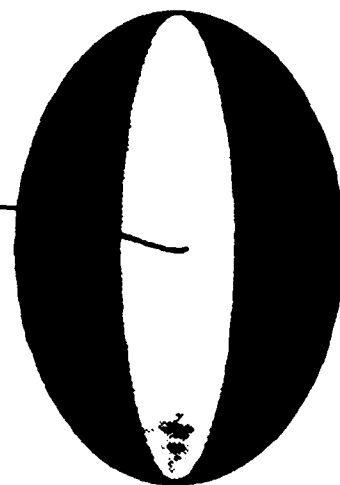
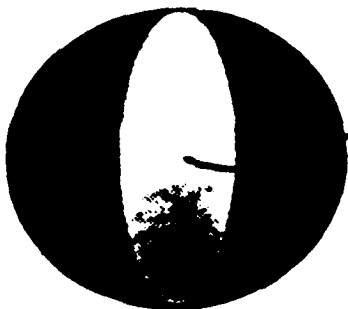
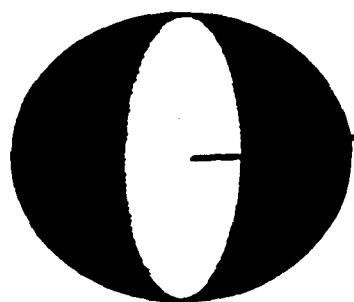
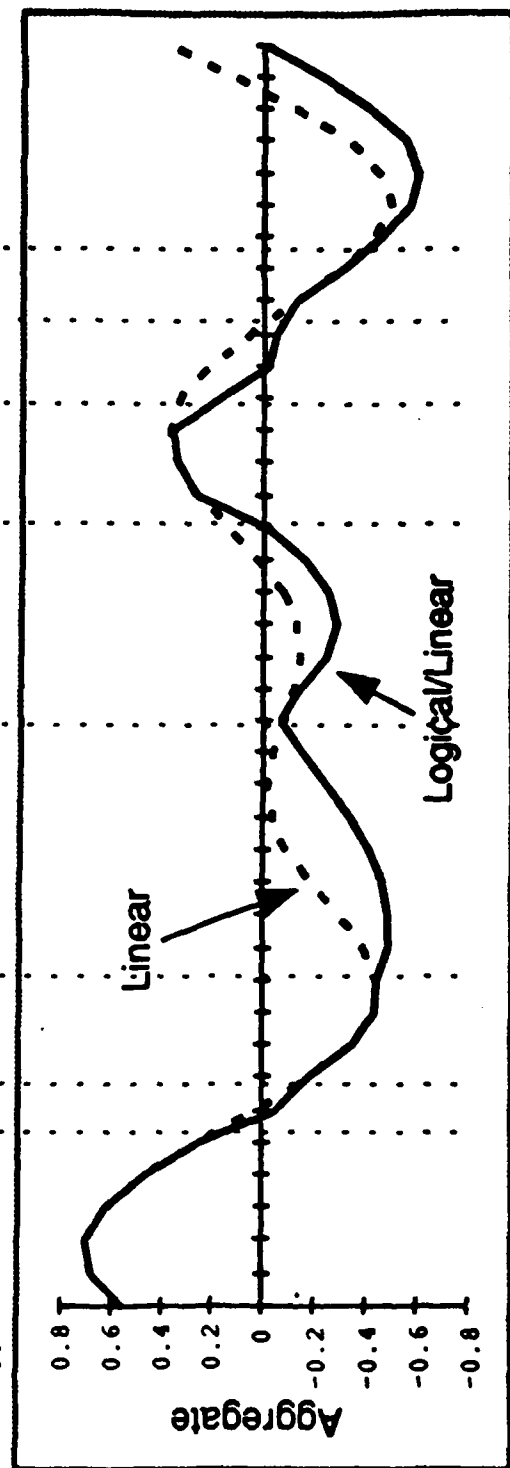
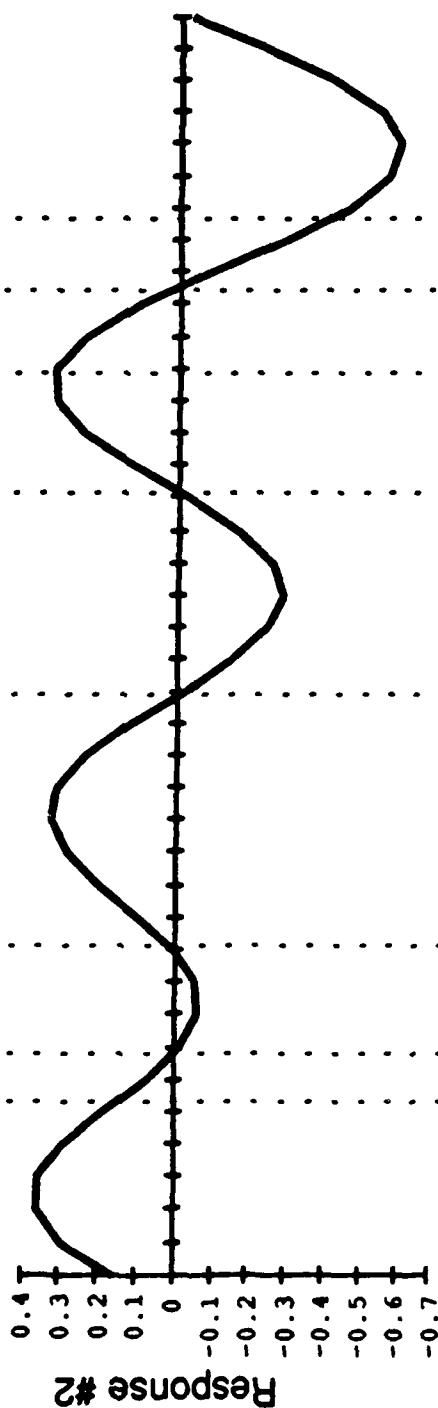
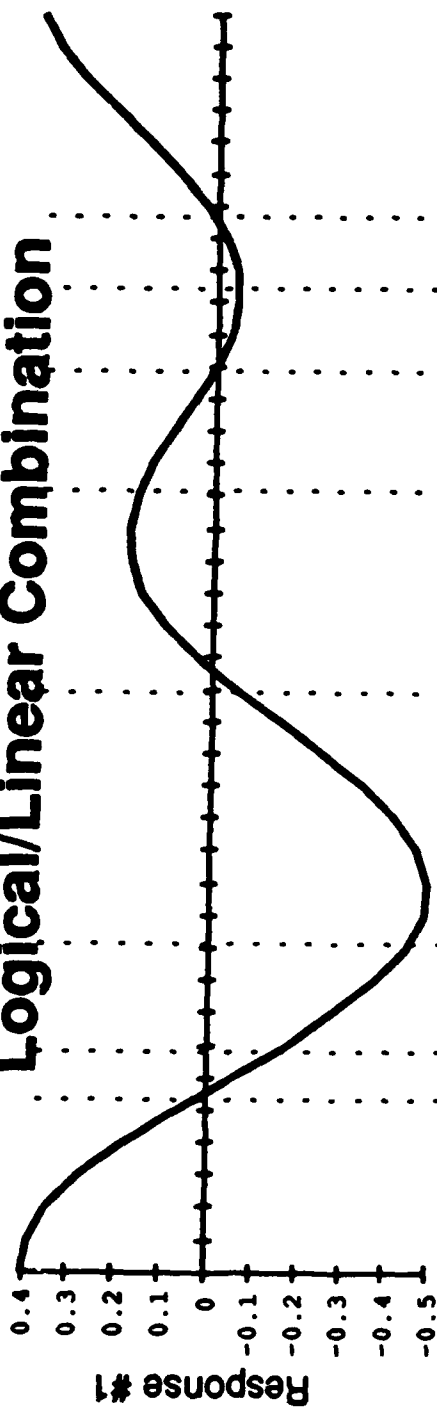
Logical/Linear Cell Construction



3. Combination of Subunit Responses

The logical/linear (L/L) combination of two subunits is illustrated for the pair of varying contrast patterns shown. The *logic* of these combinations assumes that positive responses are confirmation (**TRUE**) and negative are rejection (**FALSE**). When the outputs of both subunits are positive, the combined response is the sum of the responses. If either subunit response is negative the combined response is the sum of the negative responses. In many cases combination is followed by additional half wave rectification.

Logical/Linear Combination



4. Standard Experimental Assays

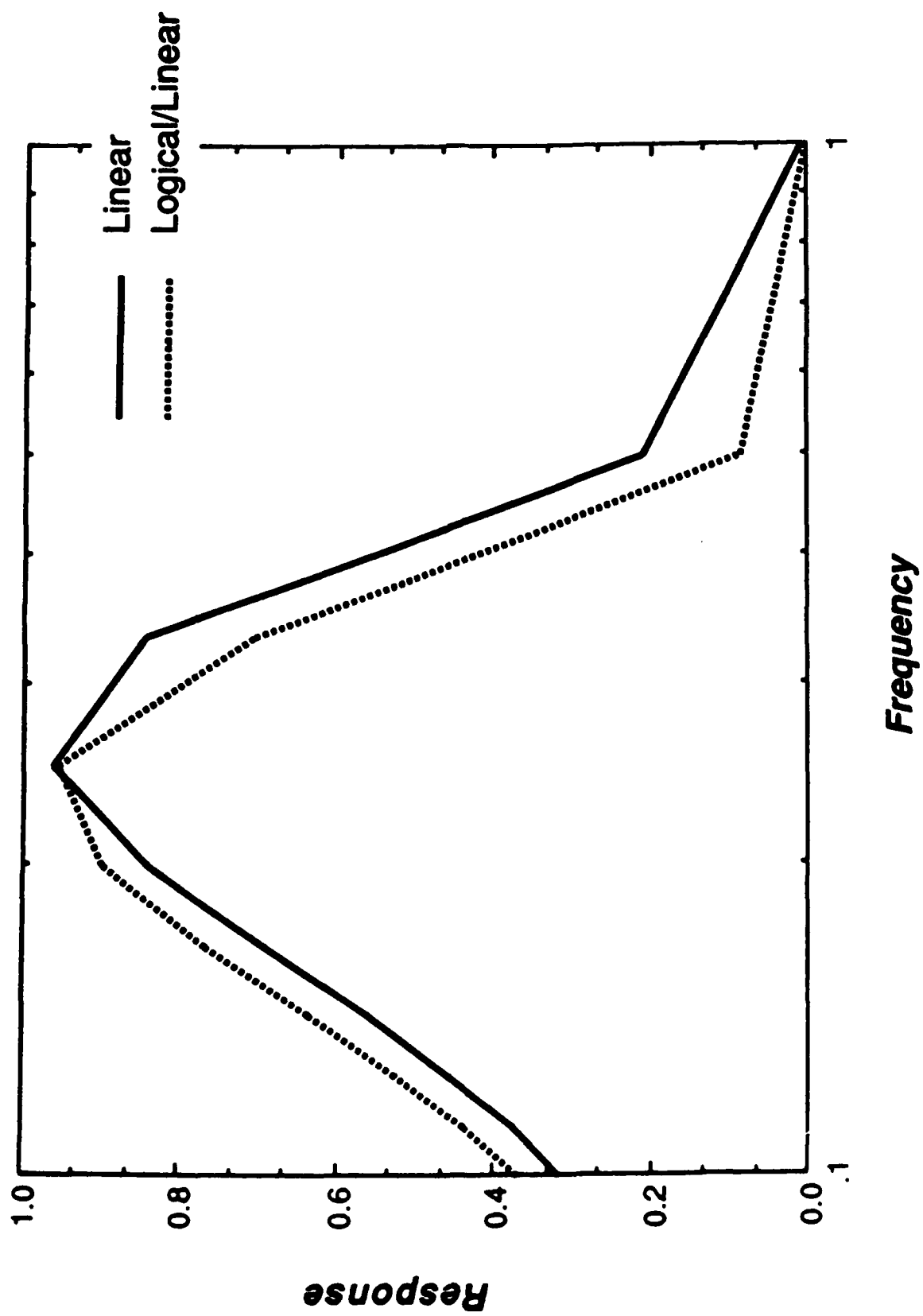
We examine the response of the linear model and the logical/linear model with bar and sinusoidal grating stimuli. The linearized RFs of all models are identical.

Frequency Response. There is essentially no difference in the magnitude of the frequency response between the linear model and its logical/linear counterparts. The spatial phase response is half wave rectified in each case. The logical/linear phase response is narrowed in a way that is indistinguishable from a thresholded linear cell.

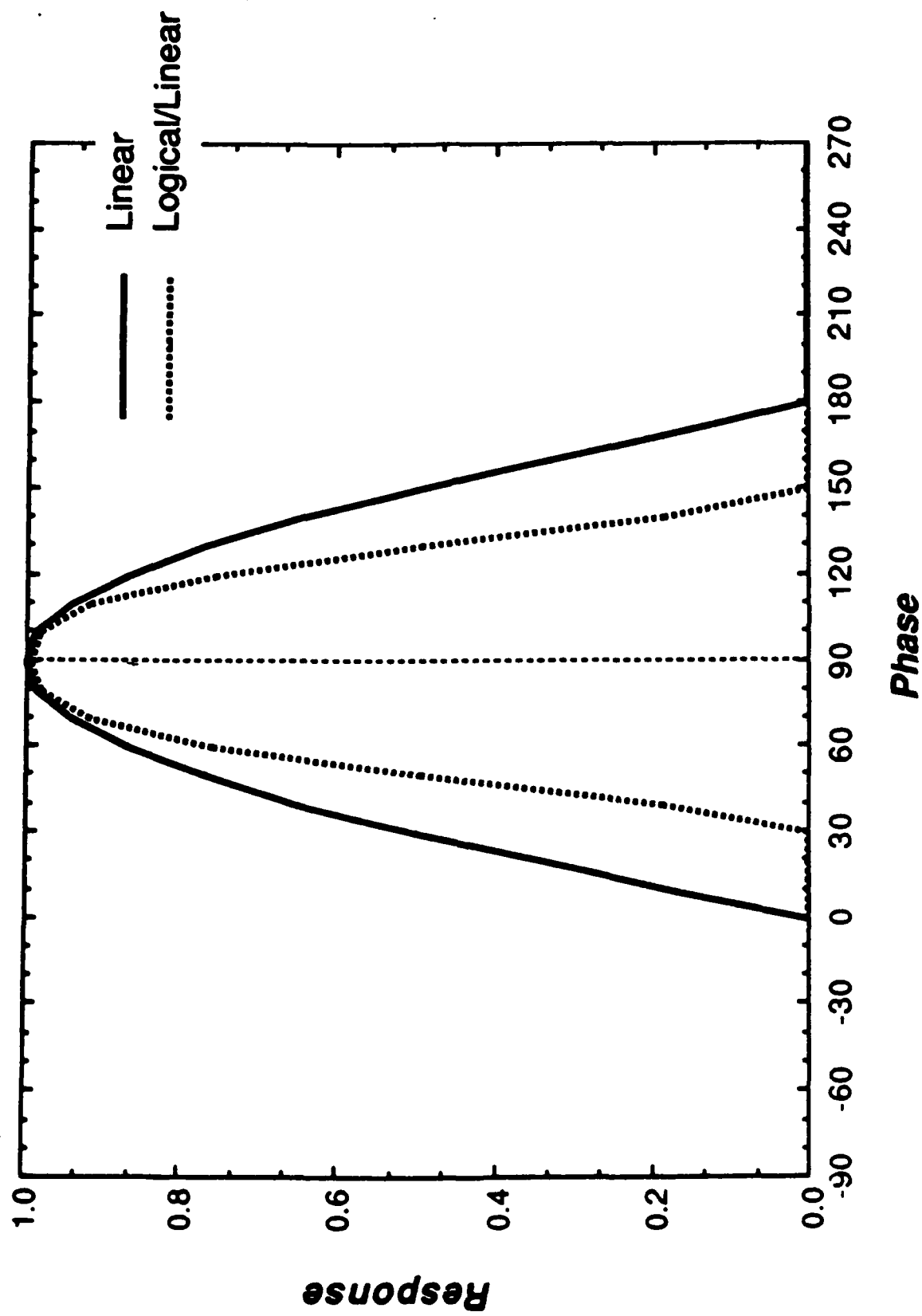
Orientation Response. The response to optimal phase and frequency sine gratings and bars is computed as a function of orientation. In each case the logical/linear model has significantly narrower tuning than the linear model.

Length Response. The response to optimal width bars is computed as a function of bar length. The logical/linear model responds less for short lengths. Although this resembles a standard response nonlinearity it is effectively a *structural* threshold — increasing the contrast would not appreciably affect the shape of the curve.

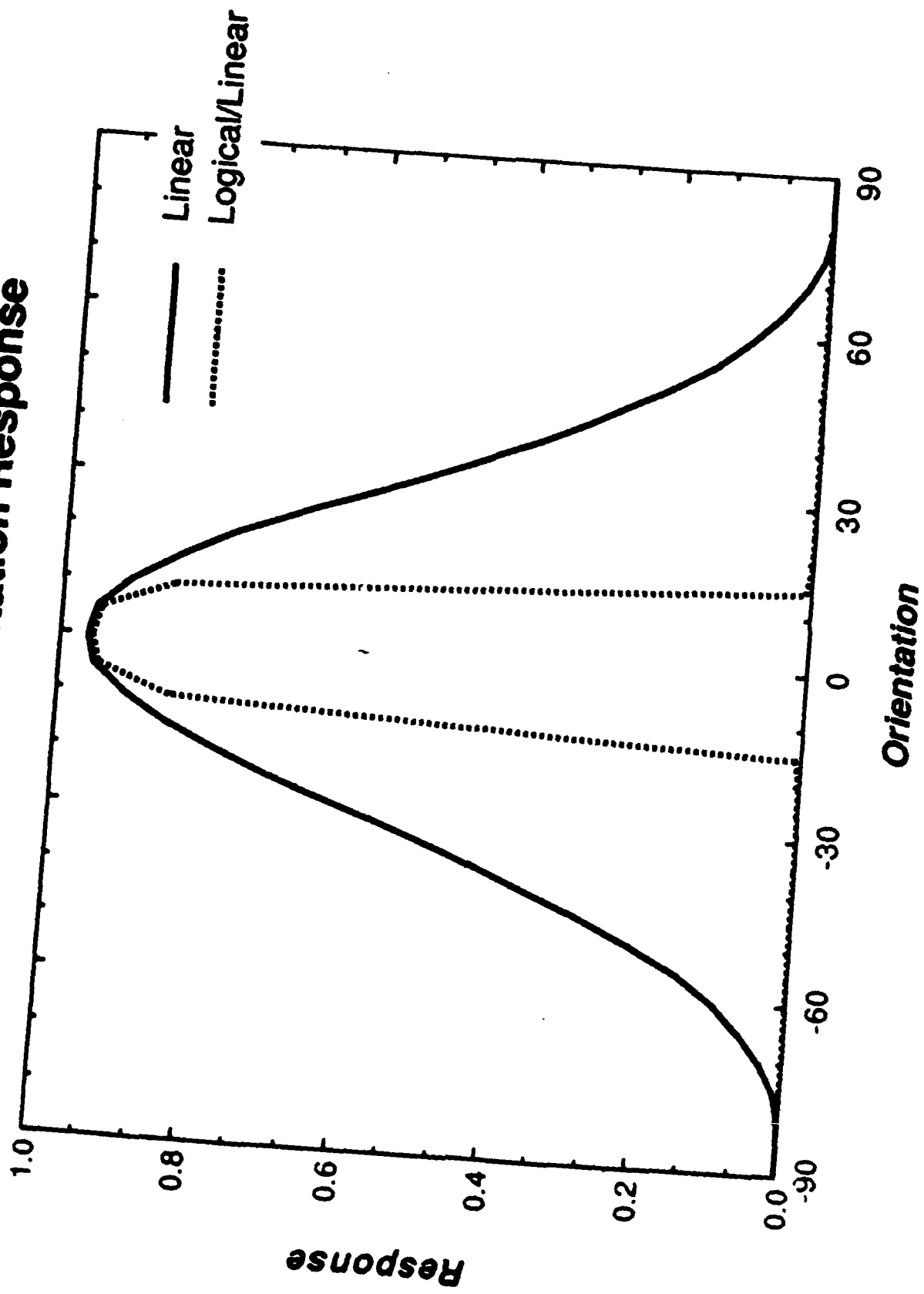
Grating: Frequency Response



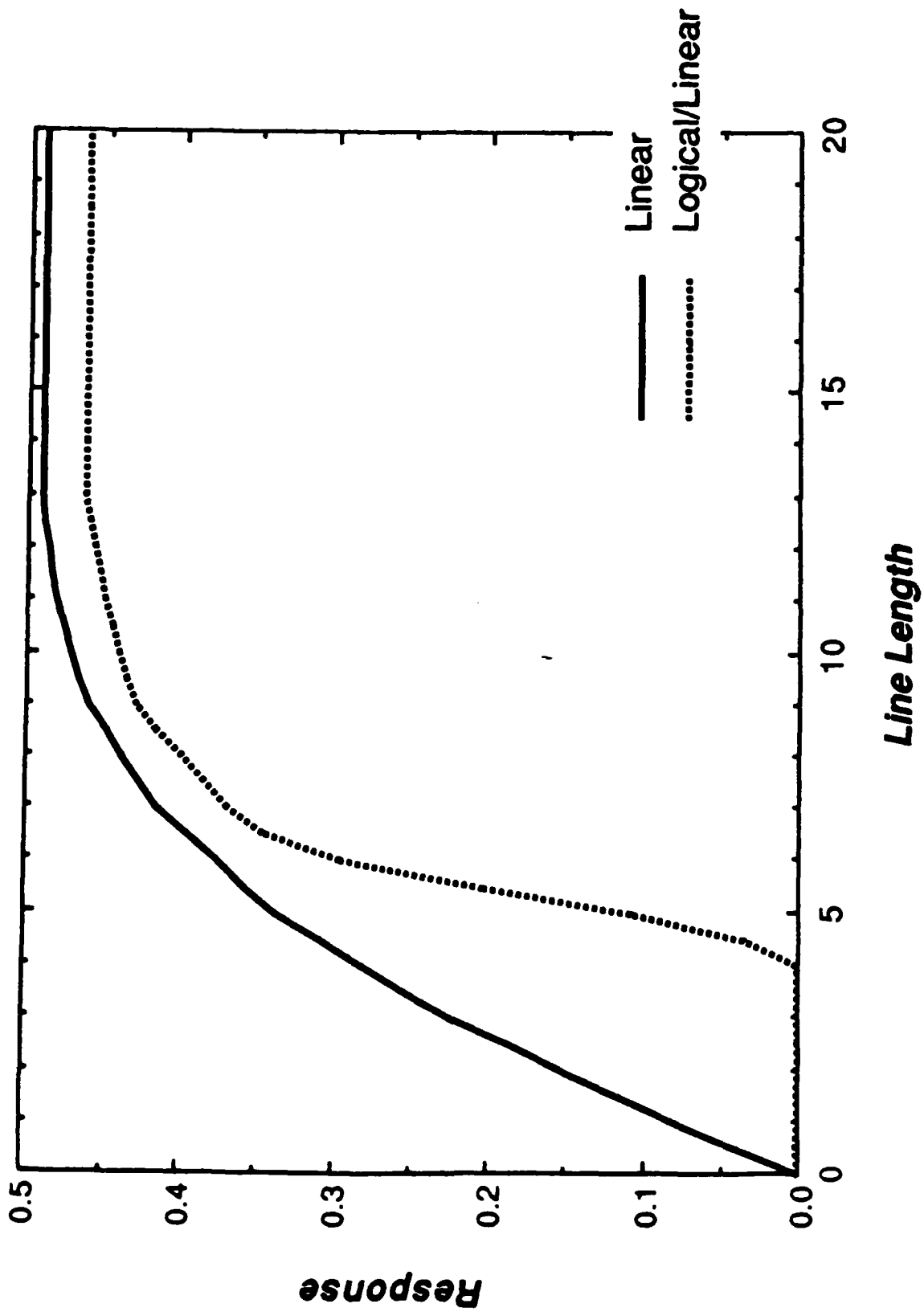
Grating: Phase Response



Grating: Orientation Response



Line: Length Response



References

- Allman, J., Miezín, F. & McGuinness, E. (1985) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Rev. Neurosci.* 8: 407-430.
- Andrews, D.P. (1967) Perception of contour orientation in the central fovea. *Vision Res.* 7: 975-1013.
- Bonds, A.B. (1989) Role of Inhibition in the Specification of Orientation Selectivity of Cells in the Cat Striate Cortex. *Visual Neurosci.* 2: 41-55.
- De Valois, K.K., De Valois, R.L., and Yund, E.W. (1979) Response of striate cortex cells to grating and checkerboard patterns. *J. Physiol. (Lond)* 291: 483-505.
- Hammond, P. & MacKay, D.M. (1983) Influence of luminance gradient reversal on simple cells in feline striate cortex. *J. Physiol. (Lond)* 337: 69-87.
- Morrone, M.C., Burr, D.C. & Maffei, L. (1982) Functional Implications of Cross-Oriented Inhibition of Cortical Visual Cells. *Proc. Roy. Soc. Lond. B* 216: 335-354.
- Swindale, N.V. & Cynader, M.S. (1986) Vernier Acuity of Neurones in Cat Visual Cortex. *Nature* 319: 591-593.
- Swindale, N.V. & Cynader, M.S. (1989) Vernier Acuties of Neurons in Area 17 of Cat Visual Cortex: Their Relation to Stimulus Length and Velocity, Orientation Selectivity, and Receptive Field Structure. *Visual Neurosci.* 2: 165-176.
- Westheimer, G. & McKee, S.P. (1977) Spatial configurations for visual hyperacuity. *Vision Res.* 17: 941-947.
- Wilson, H.R. (1986) Responses of Spatial Mechanisms Can Explain Hyperacuity. *Vision Res.* 26: 453-469.

Contrast Reversal

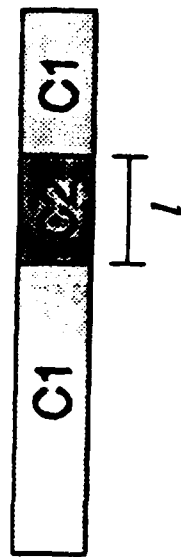
Introduction Most work on cortical cell responses to bars has concentrated on simple stimuli of consistent contrast polarity. Both De Valois, De Valois & Yund (1979) and Hammond & MacKay (1983) explored the effects of luminance gradient (contrast) reversal on simple cell responses.

Procedure The model was presented with bar stimuli at the preferred contrast and orientation. A small internal segment ('blob') was inserted in these bars, and the position and contrast of this segment was varied. This is illustrated both schematically and by example.

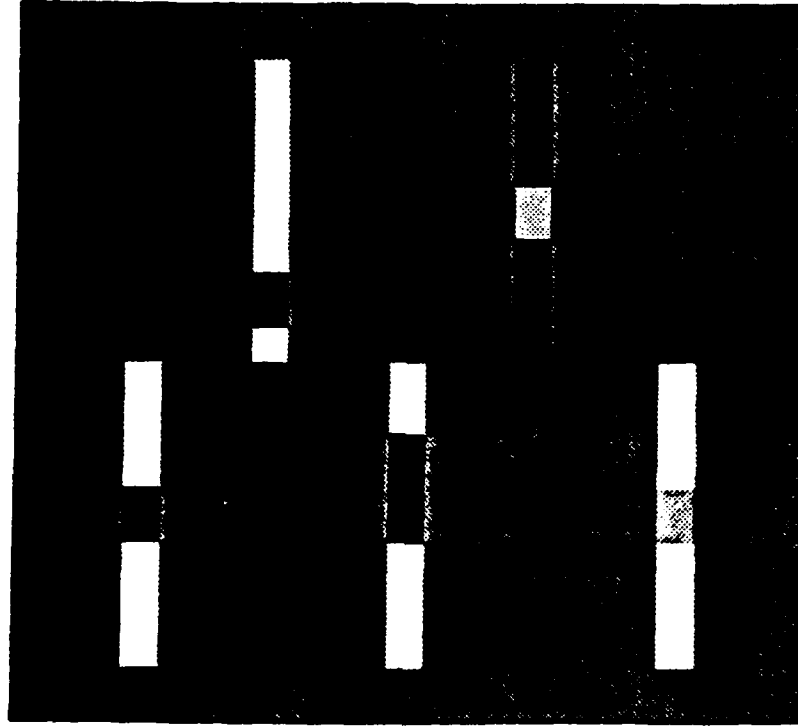
Results A substantial suppression of response is seen for the L/L model as the contrast of the 'blob' reverses polarity with respect to the base stimulus. In addition, this suppression is strongest when the segment is near the centre of the RF of the model.

Discussion The responses of the L/L model show considerable qualitative correspondence with the findings of Hammond & MacKay (1983). There is clearly a considerable non-linear suppression associated with contrast reversal of a small region in the model's RF. This effect is inherent in the model's logical combination of subunit responses.

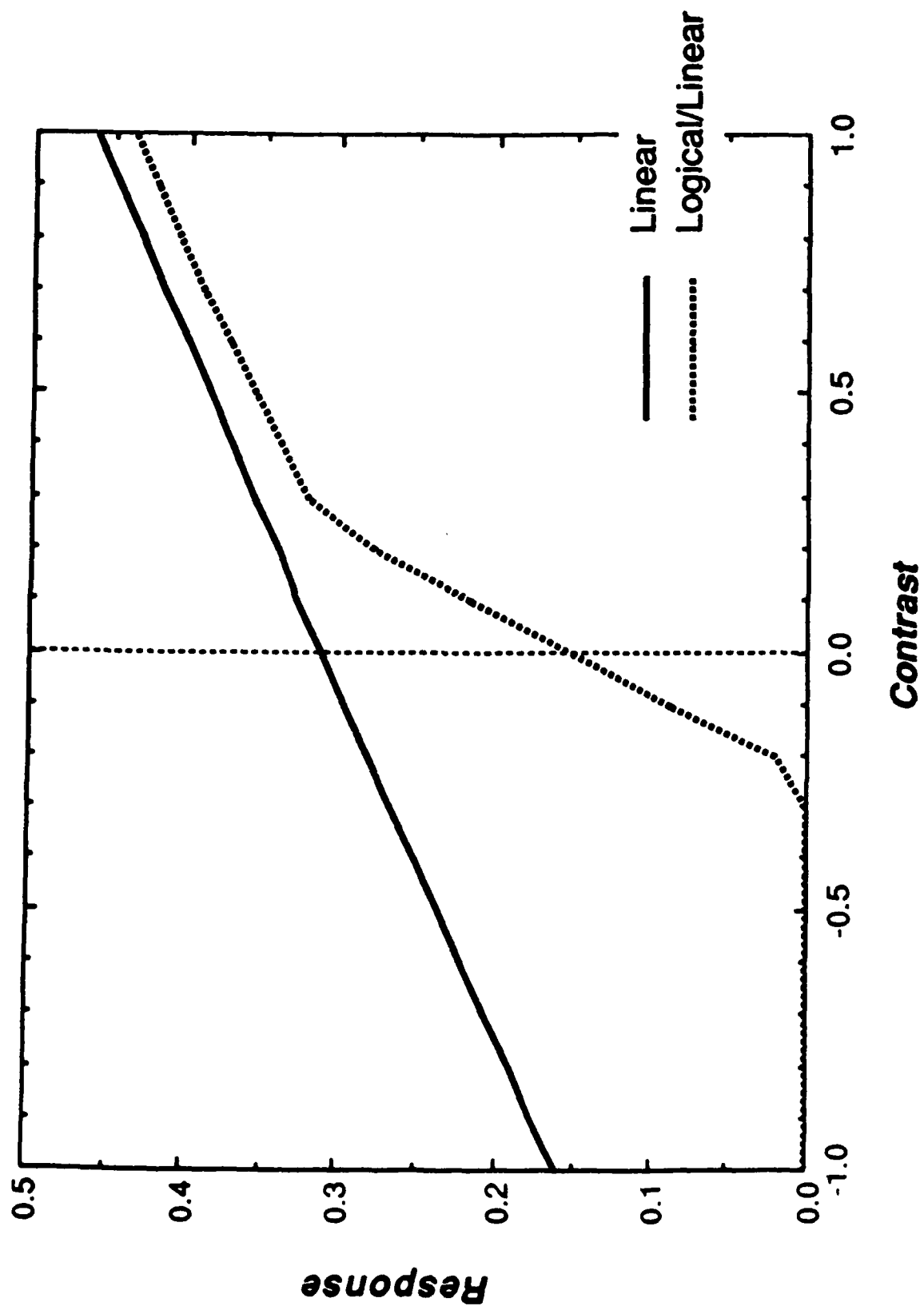
Line Contrast Variation



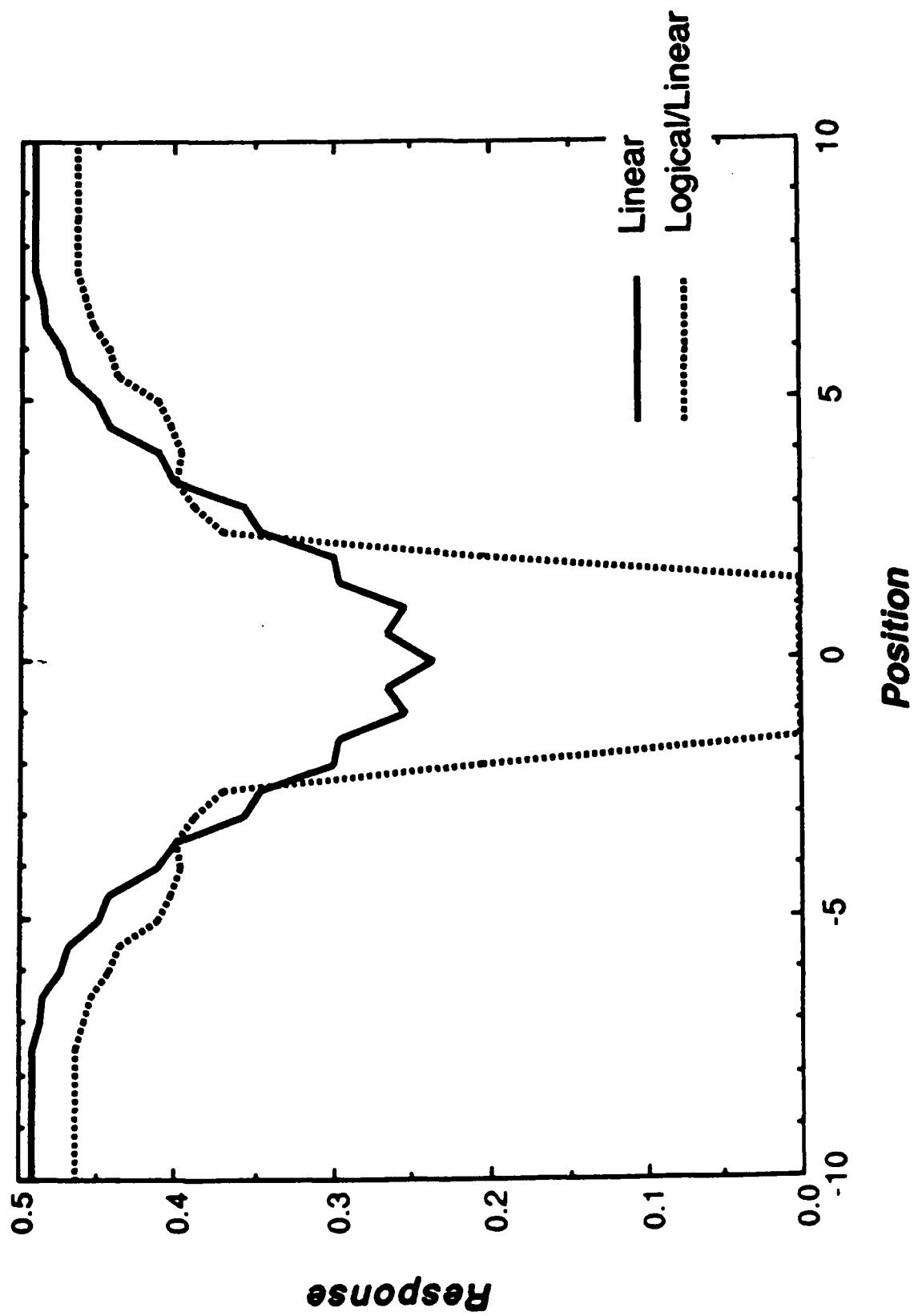
Relevant Variables:
L: Length, **W:** Width (overall),
C1: Line contrast,
C2: Blob contrast,
l: Blob length,
x: Blob position.



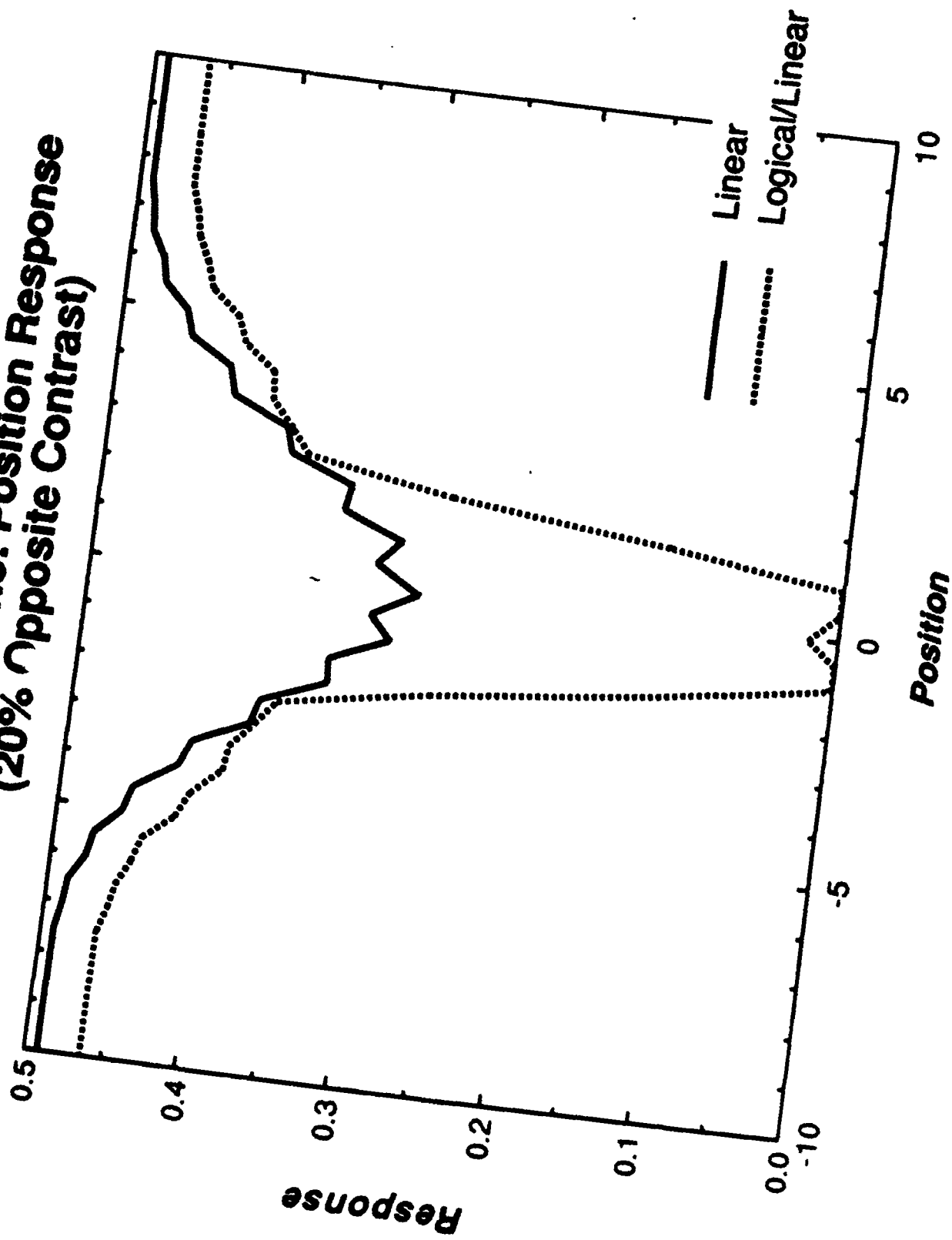
Interrupted Line: Contrast Response



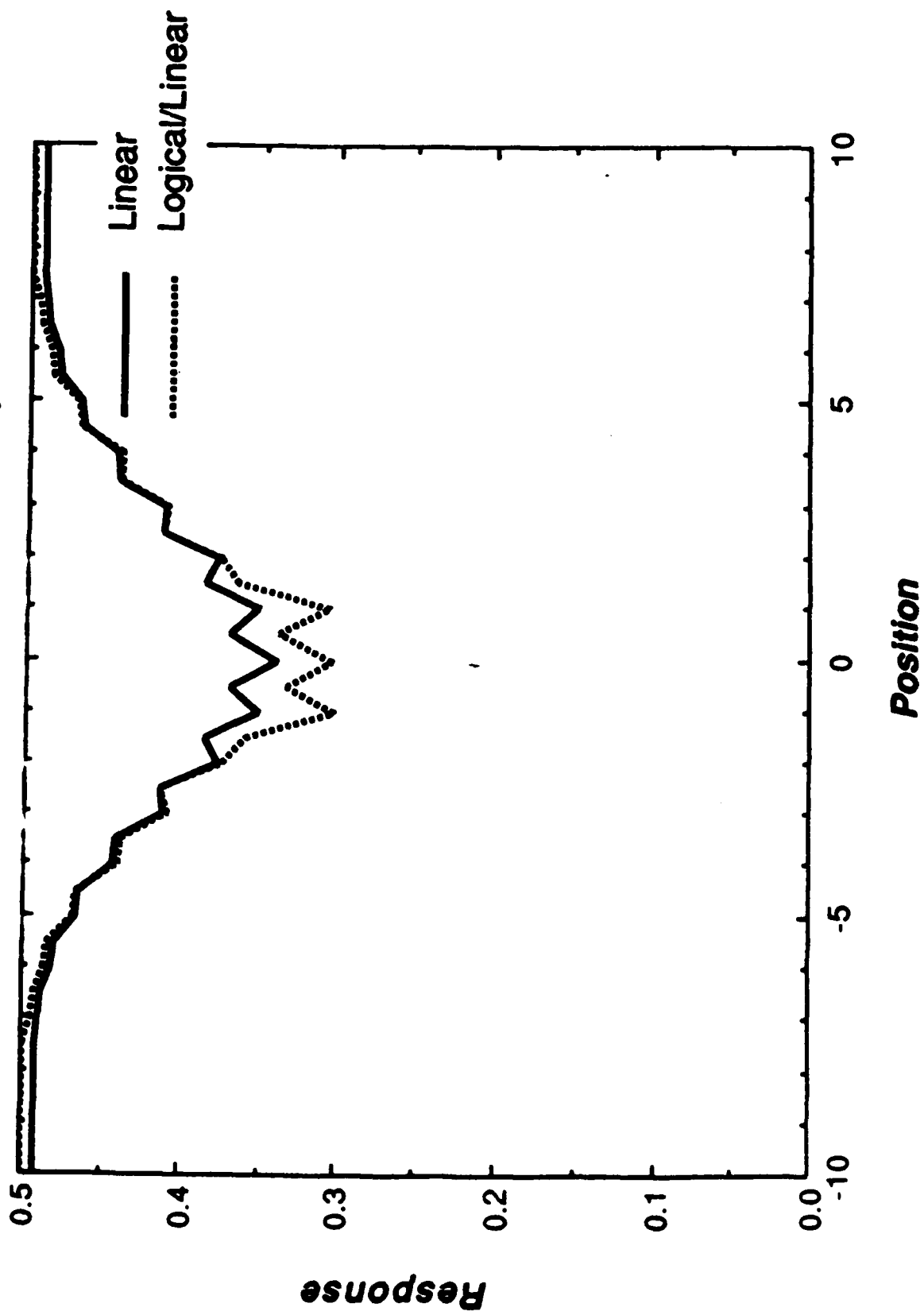
Interrupted Line: Position Response (50% Opposite Contrast)



**Interrupted Line: Position Response
(20% Opposite Contrast)**



Interrupted Line: Orientation Response (20% Similar Contrast)



Cross Orientation Inhibition

Introduction A variety of evidence supports the idea of neural interaction beyond the classical receptive field (RF). These interactions vary in specificity and spatial extent (see Allman et al., 1985 for a review).

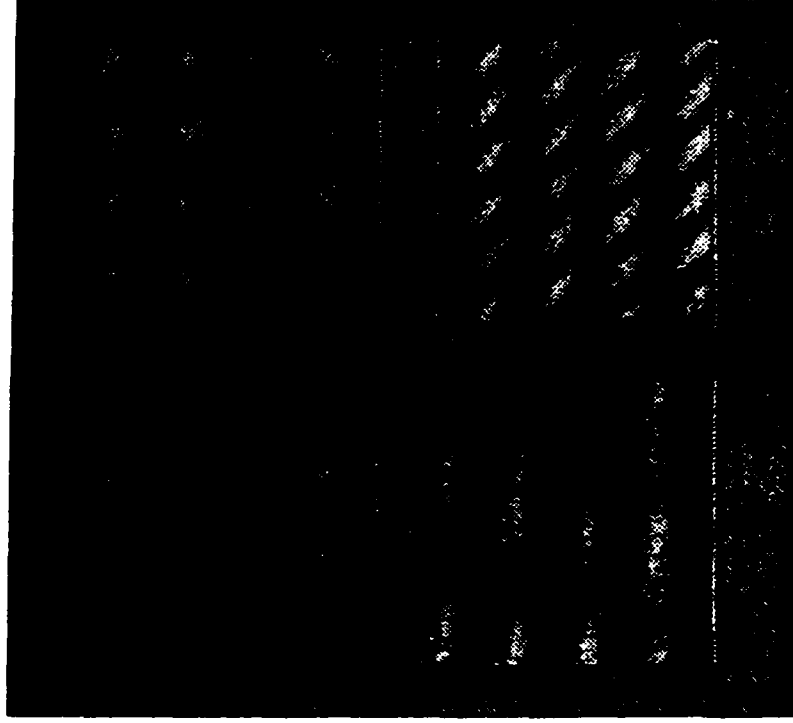
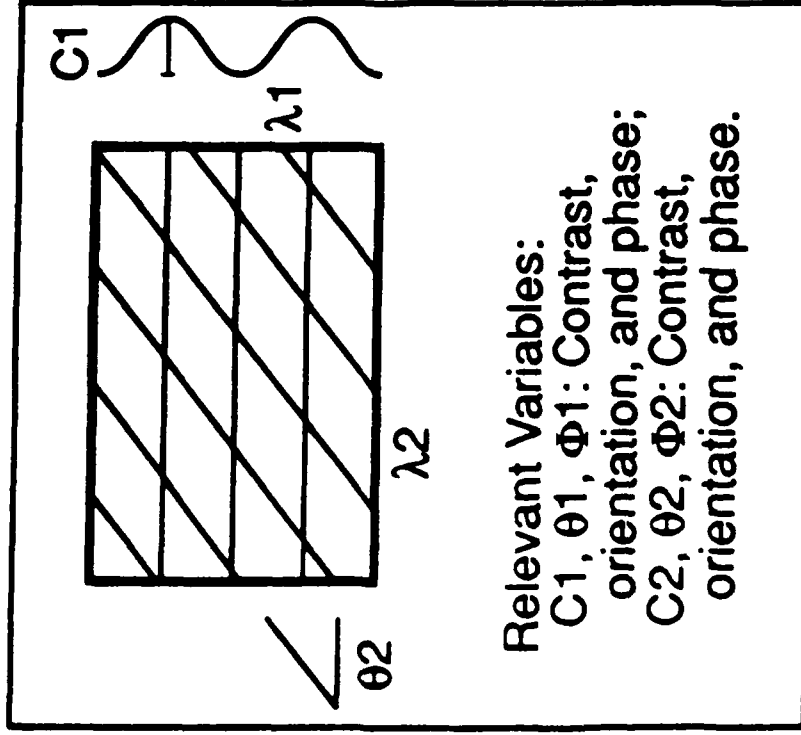
The interaction between stimuli of different orientations is frequently supposed to involve the interaction of neurons tuned to different orientations (Morrone et al., 1982). We show that *cross orientation inhibition* is an intrinsic characteristic of the L/L RF.

Procedure The response of the model to two sinusoidal gratings (or two bar stimuli) of different orientations was examined. The first stimulus was of the optimal spatial frequency and orientation while the orientation of the second stimulus was systematically varied. This is illustrated schematically in the diagram as are replicas of the stimuli.

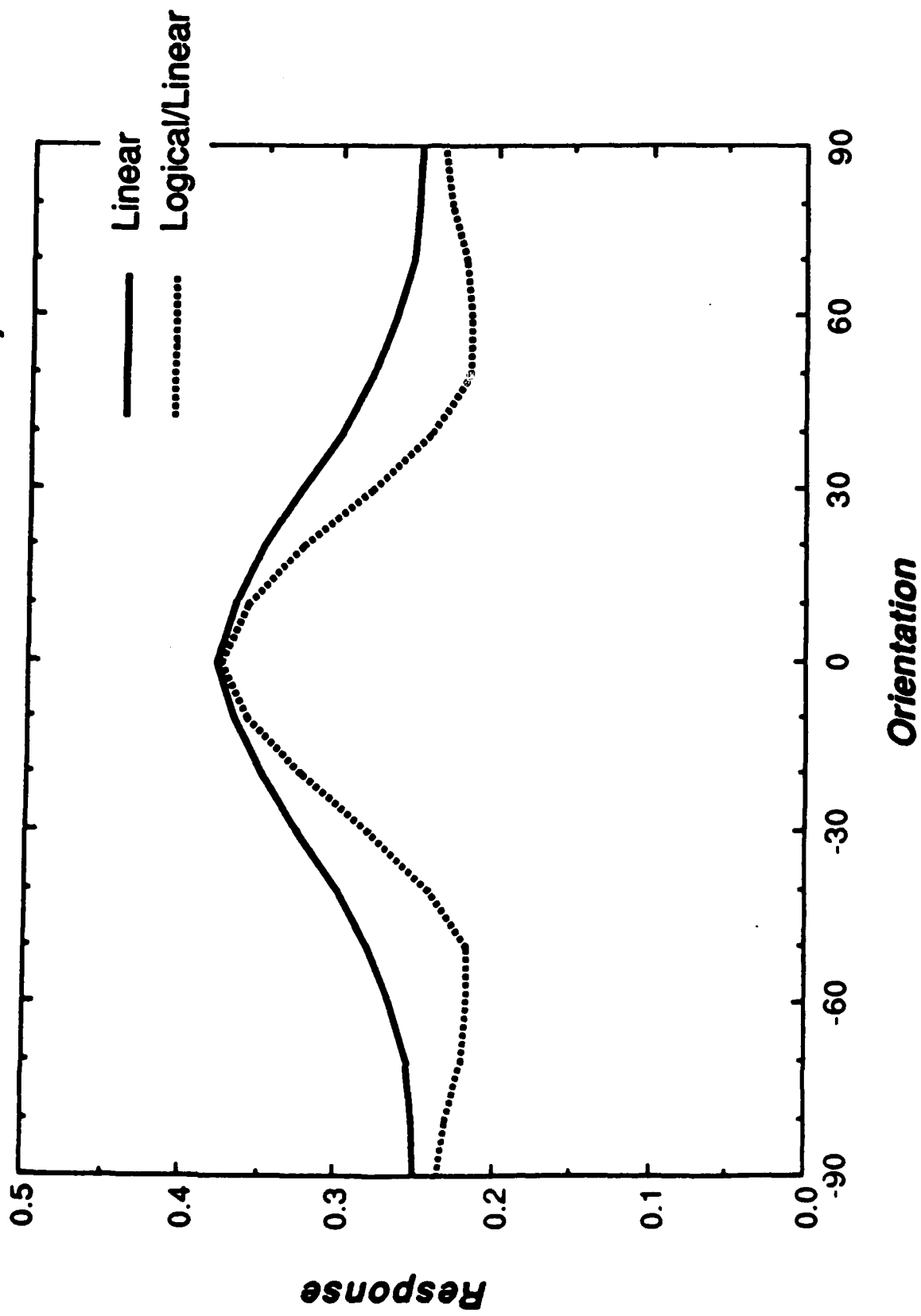
Results The solid and dashed curves show the response of instances of the linear and L/L models to the compound stimuli. For both gratings and bars, there is substantial suppression introduced by the second stimulus if it is of at least equal contrast to the base stimulus. The linear model illustrates the response attributable to spatial summation.

Discussion Morrone et al. (1982) and Bonds (1989) find the suppression introduced by second grating to be very broad — essentially uniform outside the preferred orientation range of the cell. We propose that the suppression observed is not principally an orientation effect, but may be due to the expression of reverse contrast nonlinearities in the RF.

Cross-Orientation Inhibition

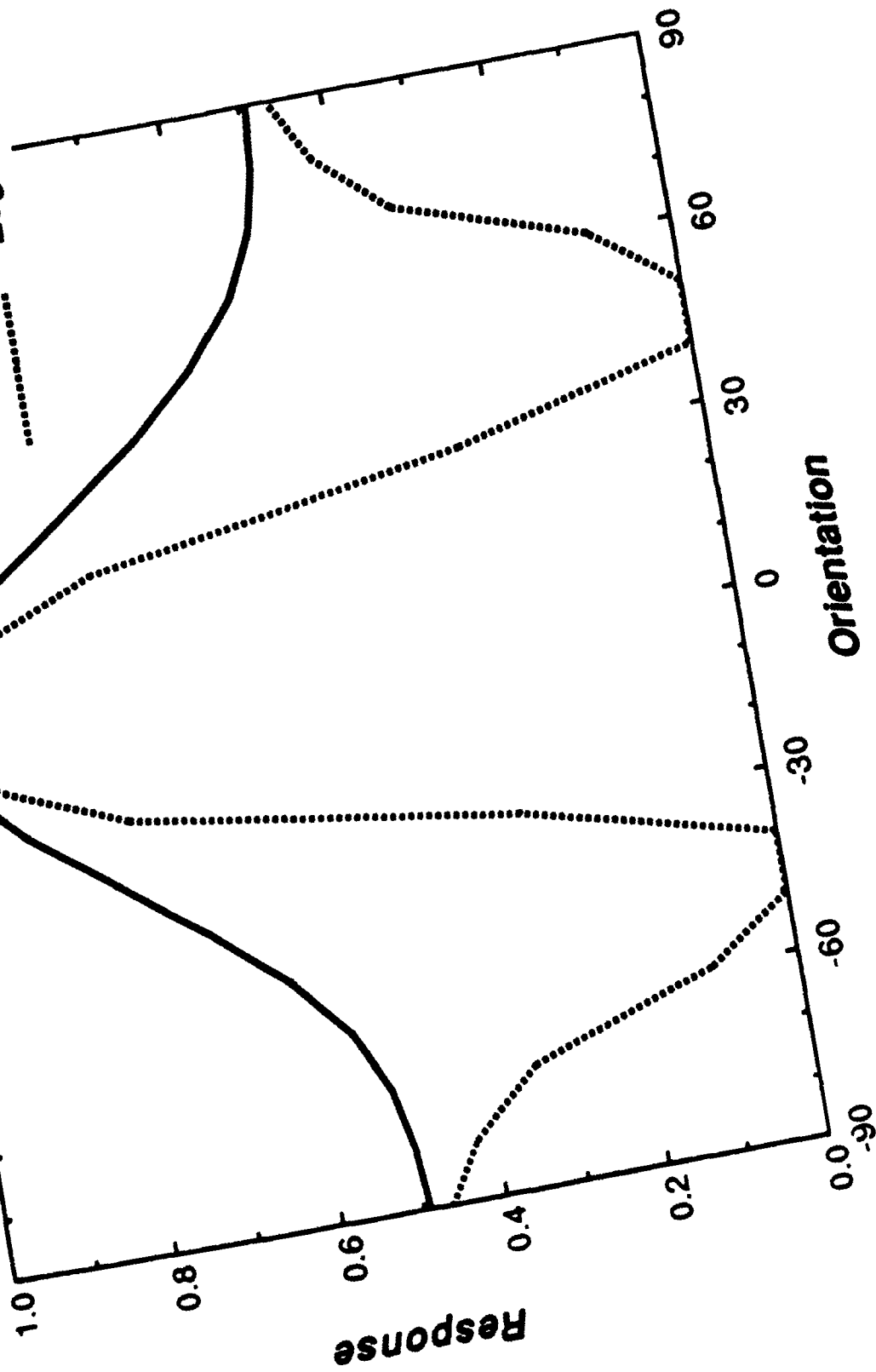


Two Gratings: Orientation Response (2nd Grating at 50% Contrast)

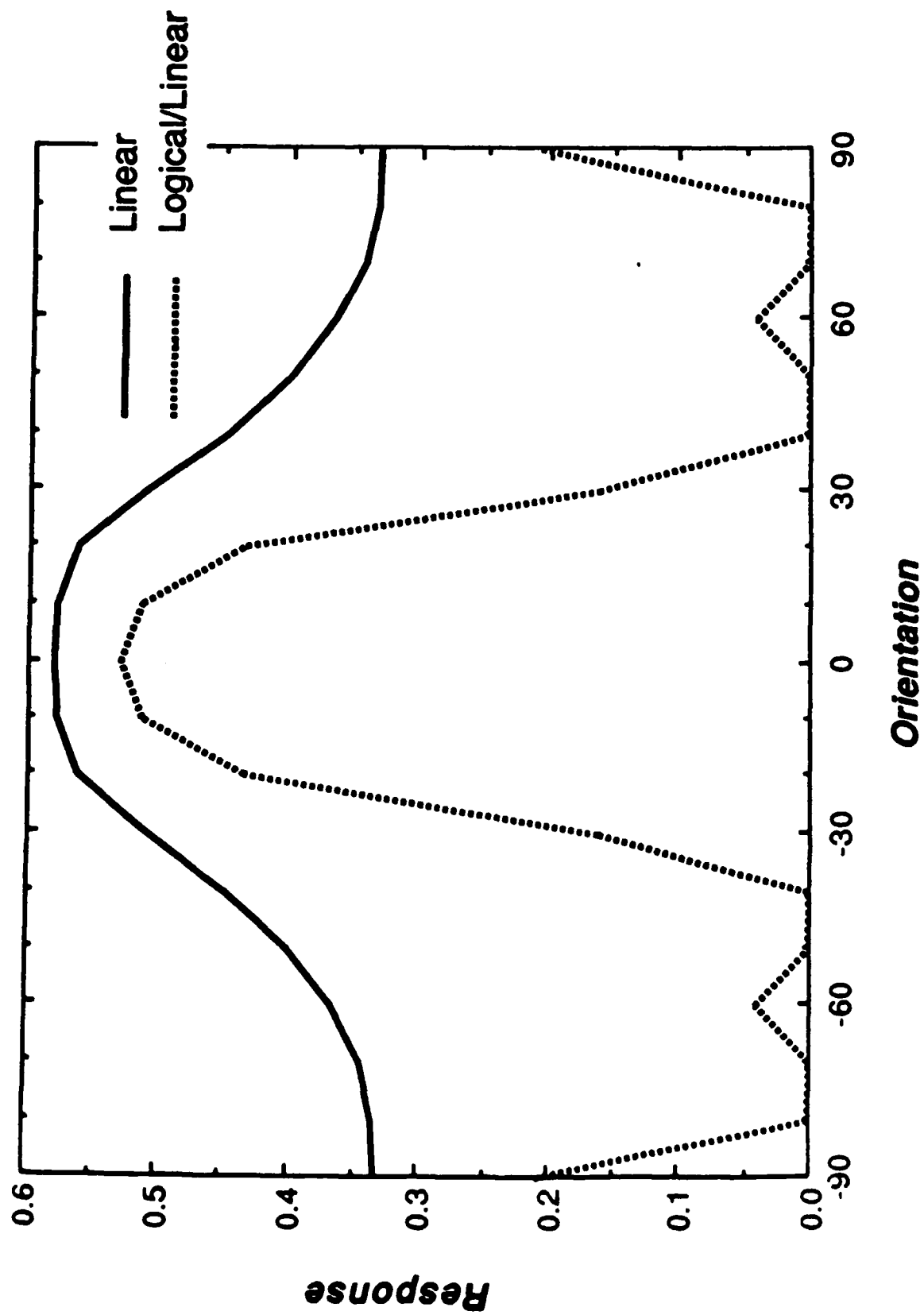


Two Gratings: Orientation Response (2nd Grating at 100% Contrast)

Linear
Logical/Linear



Two Lines: Orientation Response



Vernier Sensitivity

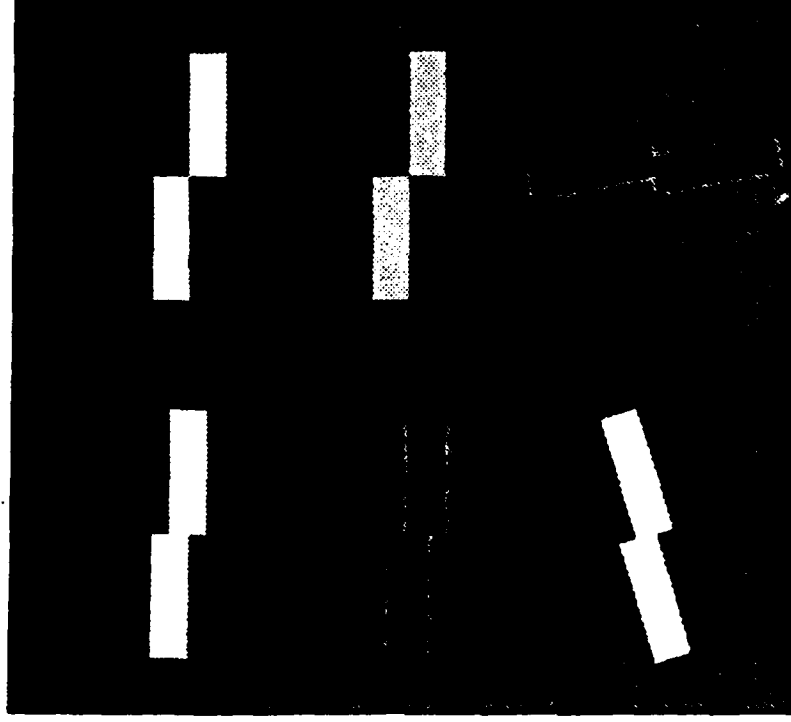
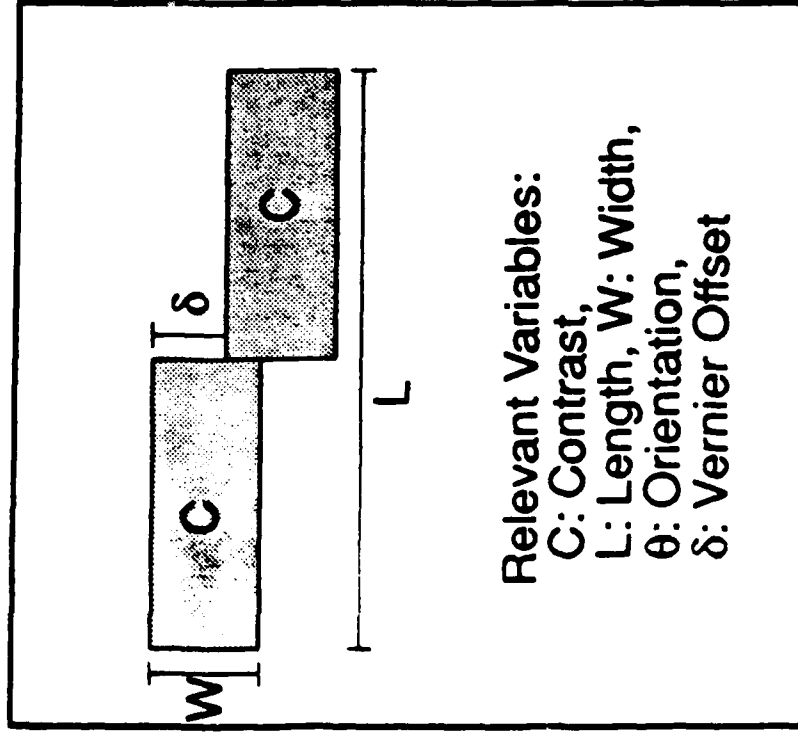
Introduction The human visual system is exquisitely sensitive to the departure from collinearity (vernier offset) of parallel line segments (Westheimer & McKee, 1977). It has been suggested that human sensitivity to vernier offsets can be accounted for by the orientation sensitivity of the most narrowly tuned cortical units (Andrews, 1967; Swindale & Cynader, 1986, 1989).

Procedure We examined the sensitivity of the linear and L/L simple RF models to lines with vernier offsets. The two bars were symmetrically displaced from the line through the RF center.

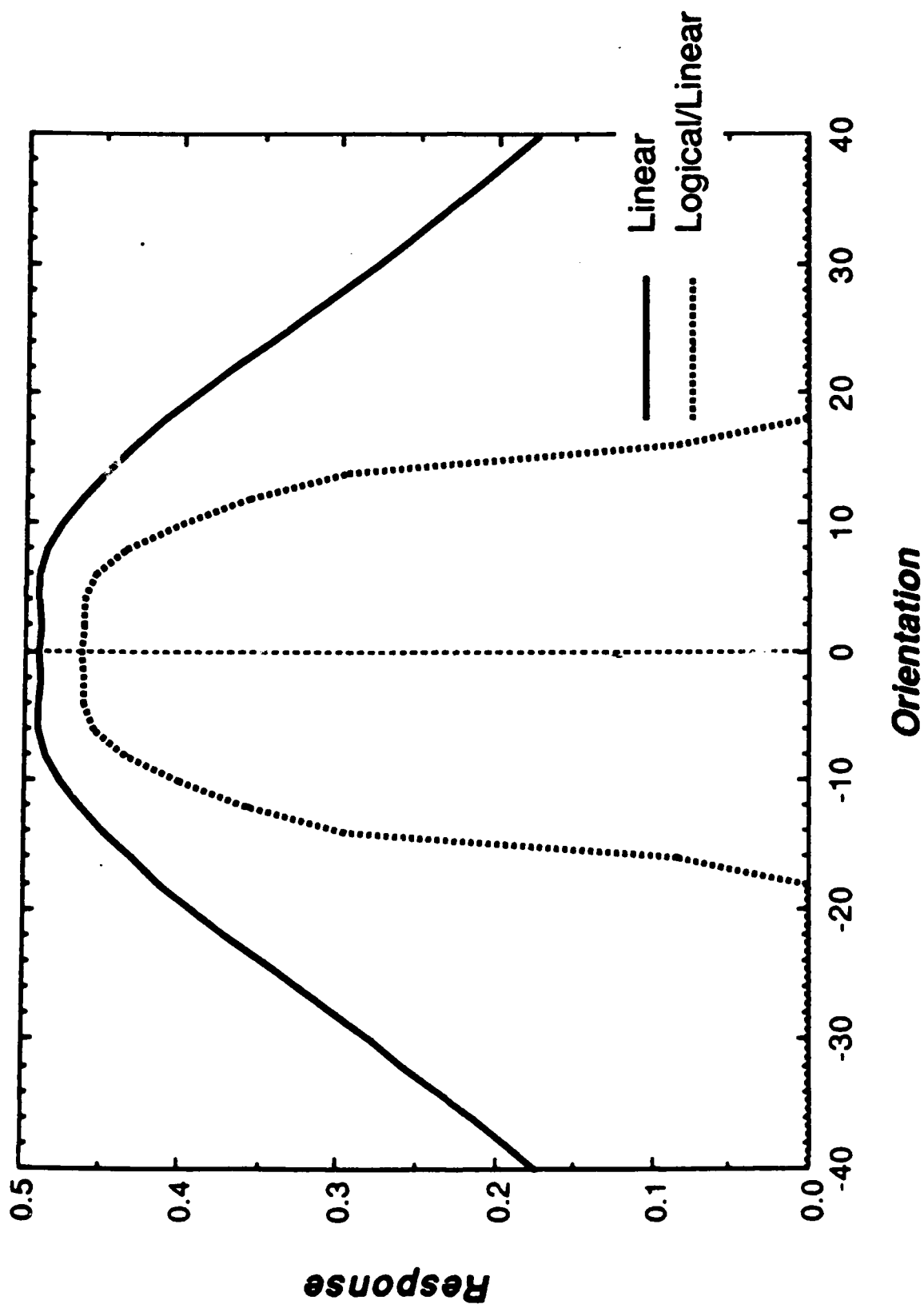
Results The solid, dotted and dashed curves show the response of instances of the linear, L/L, and stabilized L/L models to the vernier stimuli.

Discussion Swindale and Cynader (1989) found that a linear RF with threshold nonlinearity model of simple cells qualitatively accounted for the measured vernier sensitivities of cat striate cortex units. However, the measured vernier sensitivities were generally higher than those predicted by the thresholded linear model. Note that the L/L model has both narrower orientation tuning and greater vernier sensitivity than its corresponding linear RF. The combination of more local evidence (Wilson, 1986) might also contribute, but may not be necessary in light of this significant enhancement of orientation specificity.

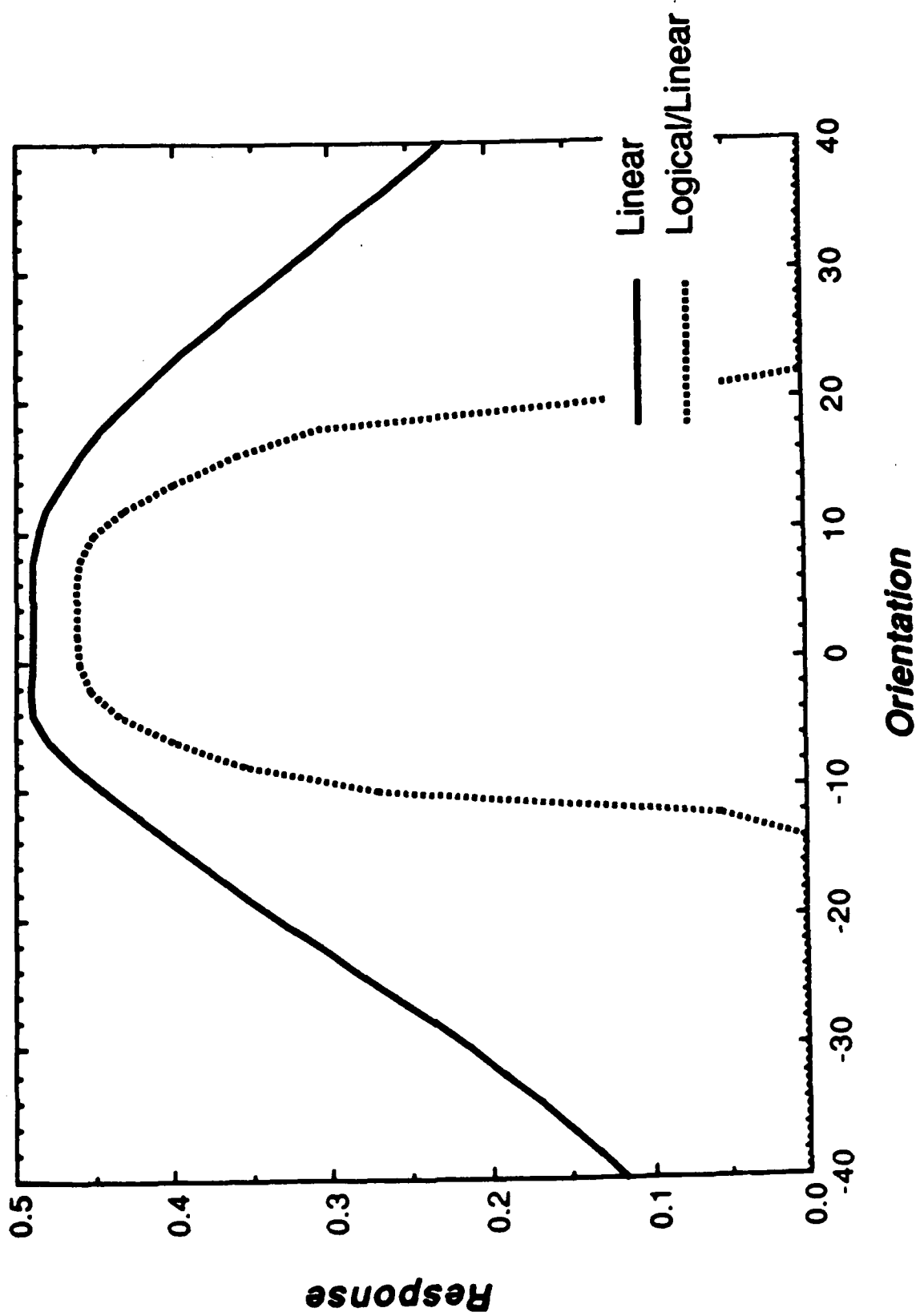
Vernier Offset



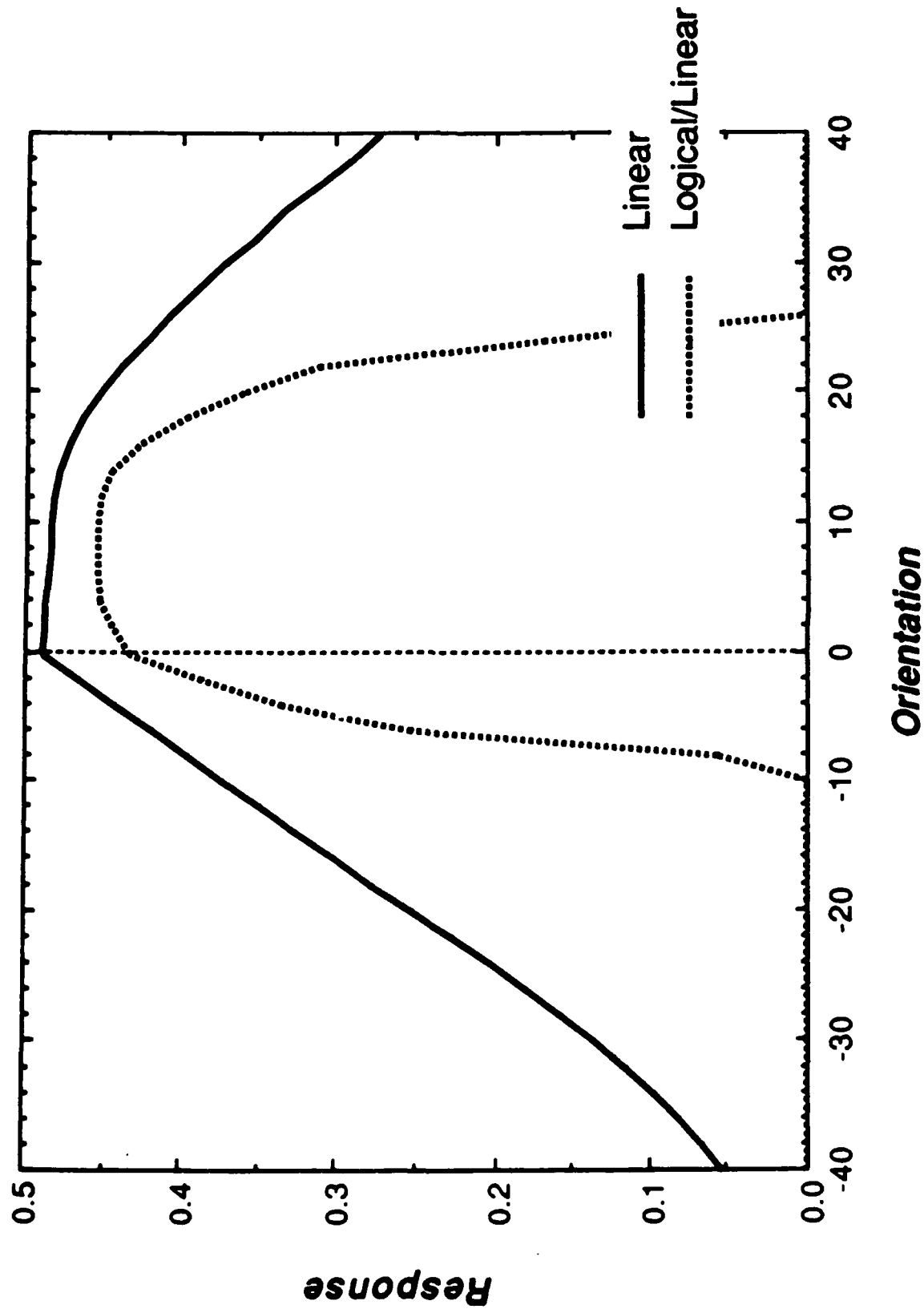
Vernier: Orientation Response (Offset 0.0)



Vernier: Orientation Response (Offset 0.5)



Vernier: Orientation Response (Offset 1.0)



Appendix III

Endstopped Visual Cortical Neurons I: Analysis of a Quasi-Linear Model.

Endstopped Visual Cortical Neurons I. Analysis of a Quasi-Linear Model.

Allan Dobbins Steven W. Zucker * Max S. Cynader †

Computer Vision and Robotics Laboratory
McGill Research Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada

* Fellow, Canadian Institute for Advanced Research

*

† Dept. of Ophthalmology, University of British Columbia, Vancouver, B.C. Canada

Running Title: Analysis of a Quasi-Linear Endstopped Neuron Model

33 pages 10 figures 0 tables

Postal Address: 3480 University Street, Montréal, Québec, Canada H3A 2A7

Telephone: (514) 398 6319 Telex: 05 268510 FAX: (514) 398 7348

Network Address: mrcim@larry.mrcim.mcgill.edu

Endstopped Visual Cortical Neurons I. Analysis of a Quasi-Linear Model.

Allan Dobbins Steven W. Zucker Max S. Cynader

Abstract

A model of endstopped simple neurons is mathematically analysed and studied quantitatively in simulation to examine its ability to provide robust local curvature estimates. The model arises from differences in the response of cells with small and large receptive fields (RFs). Conceptually, the small RF cell forms the discharge region and the large RF cell forms the inhibitory end zones. The combination of the component responses is nonlinear although under some circumstances an endstopped model instance with simple components is linear. The analysis of this simplified case provides insight into the behaviour of the significantly more nonlinear versions of the model. To study robustness, the model is examined in simulation with curves of varying radius, width and orientation. The sensitivity to perturbations of these curve parameters depends critically on the choice of model parameters. If the central discharge region and end zones are matched in spatial frequency response, then the form of curvature response is invariant along this dimension. Similarly, matching the components in orientation-tuning leads to stable curvature response when the curve orientation is perturbed. The model predicts that the response of endstopped neurons in curvature-arc length space is dominated by change in arc length for short arcs and by change in curvature for long arcs.

Acknowledgements

We wish to thank Lee Iverson, Benjamin Kimia and Peter Whaite for valuable discussion and programming support, and Patrick Cavanagh for helpful comments on the manuscript. Research supported by MRC grant MT10617 to the second author, and by NSERC grant A9939 to the third author, and by AFOSR grant 89-0260.

Endstopped Visual Cortical Neurons I. Analysis of a Quasi-Linear Model.

Allan Dobbins Steven W. Zucker Max S. Cynader

Abstract

A model of endstopped simple neurons is mathematically analysed and studied quantitatively in simulation to examine their ability to provide robust local curvature estimates. The model arises from differences in the response of cells with small and large receptive fields (RFs). Conceptually, the small RF cell forms the discharge region and the large RF cell forms the inhibitory end zones. The combination of the component responses is nonlinear although under some circumstances an endstopped model instance with simple components is linear. The analysis of this simplified case provides insight into the behaviour of the significantly more nonlinear versions of the model. To study robustness, the model is examined in simulation with curves of varying radius, width and orientation. The sensitivity to perturbations of these curve parameters depends critically on the choice of model parameters. If the central discharge region and end zones are matched in spatial frequency response, then the form of curvature response is invariant along this dimension. Similarly, matching the components in orientation-tuning leads to stable curvature response when the curve orientation is perturbed. The model predicts that the response of endstopped neurons in curvature-arc length space is dominated by change in arc length for short arcs and by change in curvature for long arcs.

Acknowledgements

We wish to thank Lee Iverson, Benjamin Kimia and Peter Whaite for valuable discussion and programming support, and Patrick Cavanagh for helpful comments on the manuscript. Research supported by MRC grant MT10617 to the second author, and by NSERC grant A9939 to the third author, and by AFOSR grant 89-0260.

1. Introduction

Endstopped neurons in the visual cortex are defined by their length-selectivity, but also differ from nonendstopped neurons in their response to curved stimuli. Hubel and Wiesel (1965) first identified hypercomplex or endstopped cells and reported that they respond well to stimuli less than some length, to corners, and to line terminations. They also provocatively suggested that endstopped cells should respond well to highly curved stimuli and might in a sense measure curvature. We have provided evidence that endstopped neurons are curvature-selective when evaluated with ideal, smooth curves, and have developed a model of endstopped simple (ES) cells which exhibits this property (Dobbins et al., 1987). The evidence and model support a functional division in which endstopped neurons respond to high curvature and endfree neurons to low curvature, each at their preferred orientation.

In this paper our specific motivations are as follows. First, to quantitatively examine the response characteristics of the quasilinear ES model (linear stages followed by rectification) and its linear approximation. The latter can be analysed in a rather complete way. Second, to examine the range of stimulus conditions under which curvature response is qualitatively invariant, and to examine *curvature resolution* by measuring quantitative response variation. Such results are indispensable to the question of the sufficiency or shortcomings of local estimates in the curve inference process. In a companion paper complex components are introduced and endstopped model instances with simple and complex components are compared.

The models considered can be summarized by a single functional form (see Eq. 1), in which the response arises from the "difference" of the responses of cells with short and long receptive fields. This characteristic can be implemented by a local circuit within an orientation column (McGuire et al., 1984; Bolz & Gilbert, 1986), and may be the simplest generic mechanism for building complex properties from simpler ones. In the endstopped simple (ES) model linear simple cell components are combined nonlinearly, while subsequent models contain one or more nonlinear

(complex) components (Dobbins et al., 1990). How the choice of parameter values — the sizes, gains and receptive field fine structure — affects the behaviour of the model is evaluated in each case.



Figure 1 About Here

Figure 1 illustrates the principal components of the model. The mapping from the input to the output of a neuron is separated into two parts: a spatial integration followed by a rectifying function. To evaluate the behaviour of the model, we systematically examine subsets of it, for example by eliminating or simplifying the nonlinearities. By this method one can see how each element contributes to the response specificity of the model.

2. Methods

Simulations were developed and run on a Symbolics 3670 LISP computer. Most of the experiments reported here examine the response of the model to curved arcs of different radii with varied curve position, orientation or width. Line and curve stimuli were either binary-valued curves with a jagged appearance, or gray-level curves (8 bit) with a smooth appearance created using a supersampling technique. In this technique the intensity value of a pixel is determined by considering the fraction of coverage of the pixel (treated as an 8 by 8 grid) by a continuous curve. Line stimuli were chosen to be less than or equal to the excitatory region in width, and (unless stated otherwise) were centered upon it.

The model is formulated in terms of integration of continuous variables over a continuous domain. The simulations, in contrast, involve summation of finite precision variables over a discrete domain. Simple cell RFs were represented as 2-D arrays of double precision floating point numbers. The array elements were the samples of the continuous RF lying on a rectangular grid.

Fine sampling was employed to minimize the difference in numerical results between rectangular-sampled and continuous RFs. That is, RFs were represented by large arrays e.g. 60 by 60 elements. Representations of Gaussians were truncated at $\pm 2\sigma$.

The simulation experiments were the following:

- (1) **Length Tuning.** Response to an optimally-oriented and positioned light bar is computed as length is added symmetrically to each end of the bar.
- (2) **Orientation Tuning.** To evaluate the orientation response of the model, a simulated light bar is positioned over the central RF and response is computed at each orientation as the bar is rotated in equal steps about its center.
- (3) **Position Tuning.** The response to bars or curves of the preferred orientation is computed as a function of position on the RF.
- (4) **Width Tuning.** The width of the bar or curve is systematically varied.
- (5) **Curvature Tuning.** Curvature response is evaluated by computing the response to curved arcs of varying arc length and radius arranged so that the mid-point of the arc falls on the central RF and oriented in such a way that the tangent at mid-arc corresponds to the preferred orientation of the RF of the cell. Except where stated otherwise the curves were semi-circular and were either thin lines or edges. In some experiments the curvature response is computed for curves rotated through some angle, or shifted in position on the RF.

3. The Endstopped Model

The essential character of the endstopped model is that its response arises from the "difference" in response of cells with receptive fields (RFs) of similar position and orientation but different size. The small (*S*) and large (*L*) component cells and endstopped simple cell (*ES*) are indicated with superscripts while subscripts (*e*, *o*) refer to even or odd RF symmetries. Each convolution result (R^S , R^L ; $R^i = k^i * I$) is passed through a function ($\phi(\cdot)$) which models the

inability of neurons to represent negative values on a low spontaneous firing baseline. Scalar constants c^S and c^L are introduced to balance the responses between S and L . Therefore the response of the endstopped cell is given by:

$$R^{ES} = \phi(c^S \cdot \phi(R^S) - c^L \cdot \phi(R^L)). \quad (1)$$

The activation (or positive part) function ϕ will be used to clarify the relationship between certain linear and nonlinear models. In this and previous papers (Dobbins et al., 1987; 1989) simple half wave rectification is employed for this purpose while in the accompanying paper a more general function is employed (Dobbins et al., 1990).

3.1 Simple Receptive Field Components

We model the response of simple neurons as spatial integrations of the RF against the image and represent the RFs by Gabor functions (see Appendix 1). This class of model has been shown to provide a good fit to the spatial RFs of simple cells (Jones & Palmer, 1987), although other models also could have served as well (Parker & Hawken, 1988). The RF parameter values are constrained to be compatible with the evidence about striate cortical simple cells. Since simple RFs have a small number of parallel zones — one or two to five (Heggelund, 1986), and the neurons do not respond to diffuse flashes, both the number of regions and the DC response component are limited. Few simple cells are exactly odd or even symmetric in their RF organization, but many are approximately so (Heggelund, 1986; Jones & Palmer, 1987). In the simulations described here only even (e) or odd (o) kernels are employed (denoted by k_e and k_o , respectively). Integration is followed by the rectification function (ϕ) described in the last section. The issues that arise in specifying the RF parameters are treated in subsequent sections.

4. Analysis of the Endstopped Model

We begin the analysis of the (non-linear) endstopped model (1) with a preliminary examination of linear RF models of endfree and endstopped cells. Although the linear model is only partly successful in capturing the characteristics of endstopped cells, it provides the foundation from which more realistic models will be developed.

4.1 Linear RF Models

The linear RFs are modeled by Gabor functions (see Appendix 1). Low modulation frequencies result in RFs like those shown schematically in Figure 2. The spectrum of their Fourier transform is also shown (a pair of Gaussians shifted by the modulation frequency with aspect ratios that are the inverse of those in the spatial domain). The simple RFs of Figure 2a (k_e (even), k_o (odd)) are modulated by a sinusoid of frequency u_o in the x direction. Introducing cosine modulation in the y direction (by v_o) yields a first approximation to endstopped RFs (Figure 2b, 2c). These are denoted (k_{ee}, k_{oe}) where the second subscript indicates the additional cosine modulation.



Figure 2 About Here

As elaborated in Appendix 1, the location of the envelopes in the frequency plane in Figure 2 provides an indication of preferred orientation in response to long lines and gratings.

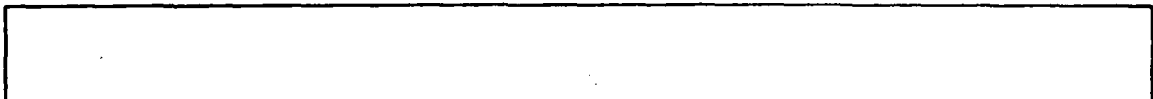


Figure 3 About Here

Let k denote a simple RF kernel and *KitsFouriertransform* :

$$k(x, y) \Leftrightarrow K(u, v).$$

The response to long lines can be found by integrating along the RF length.

$$\int_{-\infty}^{\infty} k(x, y) dy \Leftrightarrow K(u, 0). \quad (2)$$

or equivalently by examining the values of K on the u axis. Length response as a function of v_0 is shown in Figure 3a. Observe (in Figure 2b) that as v_0 increases the Gaussian envelope K (the Fourier transform of the spatial RF) moves off the u axis and therefore the response to long lines decreases. Figure 3b expresses the asymptotic end inhibition as a function of the spatial wavelength of modulation ($\lambda_y = 2\pi/v_0$). This curve, which represents the strength of end inhibition for a one parameter family of linear RFs, resembles the data on the distribution of population end inhibition as a function of discharge zone length reported for cat striate cortex (Fig. 8a in Kato et al., 1978). They found that only cells with short excitatory RFs have strong endstopping, and that long RFs invariably possess weak or no end inhibition.



Figure 4 About Here

There are three principal points concerning the orientation response of the RF types shown in Figure 2c: (i) the orientation bandwidth for short stimuli has increased; (ii) the preferred orientation for long lines is shifted toward diagonal orientations; and (iii) the orientation-tuning curve for long lines has bifurcated. Figure 4a shows that for a sequence of RFs (as in Figure 2b) the orientation response curves bifurcate as v_0 increases. Figure 4b shows the orientation response for short and long lines for a k_{ee} instance. The response to long lines is bimodal while the response to short lines is single peaked. Measured with stimuli short enough to elicit a

unimodal response curve, one finds that the breadth of response increases with increasing v_o . That is, orientation response broadens with increased endstopping. Kato et al. (1978) report that orientation tuning is broader in endstopped simple and complex cells than in their respective nonendstopped counterparts.



Figure 5 About Here

Having shown that linear "endstopped" RFs exhibit different preferred orientations for long and short stimuli we consider the response to angled stimuli. By examining the symmetries of the RFs in Figure 2c it follows that they will also respond to chevrons. The k_{ee} RF has four fold symmetry and will respond well to any of the four chevrons that can be formed from the 'X' of preferred orientations. In contrast, the k_{oe} RF responds well only to a chevron of one contrast and orientation, and its mirror in contrast and orientation. The preferred orientation of these RFs can be approximated as:

$$\theta = \pm \tan^{-1}\left(\frac{\lambda_y}{\lambda_x}\right) = \cot^{-1}\left(\frac{v_o}{u_o}\right) \quad (3)$$

Therefore the angular step ($\Delta\theta$) of the preferred chevron can be determined from the preferred orientation. In the simulation of Figure 4a the orientation peak is shifted approximately 22 degrees from vertical (Eq. 7 predicts 21.8 degrees), implying the best chevrons have interior angles of either 44 or 136 degrees. Although a chevron represents a step in orientation and an impulse in curvature, over some local neighbourhood of arc length Δs , it has the same average curvature ($\bar{\kappa}$) as some smooth curve:

$$\bar{\kappa} = \frac{1}{s_2 - s_1} \int_{s_1}^{s_2} \kappa(s) ds \approx \frac{\Delta\theta}{\Delta s} \quad (4)$$

The arc length (Δs) in the chevron curvature approximation is calculated from $\lambda_x/2$ and $\lambda_y/2$ which are approximations to local maxima of an even-symmetric RF. One can calculate the peak curvature response for $k_{\epsilon\epsilon}$ by finding the radius of the circle passing through three points (the center $(0,0)$, and $(\lambda_x/2, \pm\lambda_y/2)$):

$$\kappa = \frac{4\lambda_x}{\lambda_x^2 + \lambda_y^2}. \quad (5)$$

For smooth curves and chevrons wide enough to fill most of the central excitatory RF, and of the same average curvature, the linear RF models respond equally well, while for thinner curves there is less response to chevrons. The reason is that the smooth curve with its continuous tangent provides more excitatory drive than the corner.

Figure 5 illustrates the curvature response of the same sequence of RFs. The first two ($v_o = 0, 0.3$) respond best to straight (or low curvature) curves while the latter two ($v_o = 0.6, 0.9$) exhibit a bandpass curvature response. From (5) the predicted peak radius of curvature for $v_o = 0.6$ is 12.1 and for $v_o = 0.9$ is 6.3. From Figure 5a the peaks occur at approximately 10 and 5 respectively.

The advantage of this class of model is clearly its simplicity — at a particular RF scale (σ_x, σ_y) varying one parameter systematically varies the asymptotic end inhibition and the preferred curvature. However such models have two major disadvantages. First, from a computational perspective, since the highly endstopped instances respond well to high curvatures at one orientation and low curvatures at another, to checkerboard patterns, and to stimuli in the endzones, in isolation they can provide little information about curvature. Second, RFs with precisely these properties — bimodal orientation-tuning for long lines and excitation flanking inhibition in the endzones — have not been reported. The second of these should perhaps not be overemphasized since investigators would ordinarily take pains to avoid the endzones while assessing orientation tuning. Given these shortcomings, it is curious that the linear model explains certain observations. These

will be taken up in the discussion. The utility of the linear model derives from the fact that it can be completely analysed, and that nonlinear endstopped models reduce either exactly or approximately to a linear model in some range of operation.

4.2 A Non-Linear Endstopped Model with Simple Components

If curvature-selective responses are more than an artifact of the responses of endstopped neurons then it is necessary that the response to curves exhibit stability as other stimulus parameters change. There are at least two possible senses of stability. A strong sense is curvature-response being invariant with changes of other curve parameters. A weaker sense of stability allows qualitative change in curvature response provided it occurs at a response level significantly diminished compared to the unperturbed stimulus. To be concrete, an isotropic Laplacian of Gaussian RF exhibits curvature response that is invariant with respect to curve orientation, whereas a difference of oriented operators yields orientation-dependent curvature response. In this section we examine how the choice of model parameters affects the stability of curvature-response for model instances in which both the small and large RFs model cortical simple cells (ES model).

4.2.1 The Relationship Between the Linear and ES Models

Consider how the positive part function, ϕ connects linear RF models, like those just analysed, to the ES model (1) with linear components. If the integrations for the small and large components are in the linear range ($c^s \cdot r^s > c^l \cdot r^l$, $c^s \cdot r^s > 0$), then the combination of the two is linear. for certain choices of the parameter values the es model then closely resembles a single linear rf model (as depicted in figure 2c). in particular, if the small (s) and large (l) simple cells have identical spatial frequency (u_0) and phase (e, o) of modulation, then the spatial zeros of their rfs correspond and the difference of the two has the same center frequency (but not necessarily bandwidth). for those stimuli which bias the response of both components into the linear range,

the es model closely approximates the linear gabor model. that is, for appropriate choice of c^s and c^l , a difference of small and large even or odd simple rfs:

$$c^s \cdot k_e^s - c^l \cdot k_e^l \approx k_{ee}. \quad (6a)$$

$$c^s \cdot k_o^s - c^l \cdot k_o^l \approx k_{oe}. \quad (6b)$$

is approximated by linear endstopped rfs (k_{ee} , k_{oe}). therefore, over some range of conditions, such an es model will exhibit the same assets and liabilities as the linear models examined in the last section. viewing the linear endstopped rfs as a difference of nonendstopped rfs of different length illuminates the basis of curvature-selectivity from a different perspective. having chosen $u_o^s = u_o^l$, eq. 6b implies that s has greater orientation bandwidth than l ($bw_\theta^s > bw_\theta^l$). the difference of the two, δbw_θ , is precisely what appears in the bifurcated orientation curves of figure 4. hence, since δs is implicit in the rf scale, the average curvature (eq. 4) to which the rf is tuned can be obtained from the orientation curve by dividing the distance between the peaks by the circular arc length. the breadth of curvature tuning can also be approximated from the orientation response by measuring the distance between the proximal and distal half height points of the bifurcated response. for es instances composed in this way the orientation-tuning to long lines predicts the curvature-response.

at a particular scale the nonendstopped cells respond best to low curvatures while endstopped cells respond best to high curvatures. across scales, however, a more general statement is that endstopped cells respond to arcs that span large orientation changes and nonendstopped cells to arcs that span small orientation changes.

4.2.2 Varying Parameters Determining Spatial Frequency and Orientation Response

In order for endstopped neurons to provide reliable information about the rate of curve

bending, the discharge zones and the endzones must satisfy certain constraints. These issues are explored for neurons with linear components in this section.

Constraints between parameters derive from the Uncertainty Principle for 2D linear filters (Daugman, 1985). The Uncertainty Principle in 1D holds that the product of the square root of the variance of the modulus of a (normalised) signal in one domain ($\Delta x = [\text{Var}(k^*(x) \cdot k(x))]^{1/2}$) and the transform domain ($\Delta u = [\text{Var}(K^*(u) \cdot K(u))]^{1/2}$) is bounded below by a definite value. In this case the signal is the RF kernel. In two dimensions, two 1D relations hold as well as a 2D relation:

$$\Delta x \Delta u \geq \frac{1}{4\pi} \quad (7a)$$

$$\Delta y \Delta v \geq \frac{1}{4\pi} \quad (7b)$$

$$\Delta x \Delta u \Delta y \Delta v \geq \frac{1}{16\pi^2} \quad (7c)$$

Evidently, although the parameters of linear filters can be chosen in various ways, these relations specify limits on the characteristics that can be achieved simultaneously in both domains. Daugman (1985) analyses tradeoffs between orientation and frequency bandwidths which turns out to be significant in determining the robustness of curvature-response. It can be shown that RFs of different size cannot be simultaneously matched in spatial frequency (peak and bandwidth) and orientation bandwidth. Given the RF sizes ($\Delta y^L > \Delta y^S$) and the relation determining orientation bandwidth ($BW_\theta = \lambda_x / \Delta y$ (A6)) this result is evident.



Figure 6 About Here

Figure 6 shows different combinations of small and large RFs in the space and spatial frequency domain. In Figure 6a the Gaussian envelopes of the small and large simple RFs differ in length but not width and both are modulated by a sinusoid of the same frequency u_0 . Therefore in the frequency domain they differ only in the v dimension. In Figure 6b the Gaussian envelopes have the same aspect ratios and hence differ in width as well as length. Accordingly, in the frequency domain the envelopes differ in extent in both u and v dimensions. Figure 6c is like 10a except that the larger spatial envelope is modulated by a lower frequency than the smaller. Therefore, the smaller (spatial) envelope is shifted farther from the origin in the frequency domain. Figure 6d shows the case in which the small and large RFs differ significantly in envelope aspect ratio and hence in spatial frequency bandwidth (BW_u).



Figure 7 About Here

Figure 7a and 7b show the curvature-response of an ES instance (composed as in Figure 6a) to curved lines of different radius and width at the optimal position and orientation. When plotted on logarithmic axes it is clear that the shape of the response does not change over a range of line widths (and hence spatial frequencies). Figure 7c shows the response to curved lines of varying curvature and orientation. The response function changes both in magnitude and shape with orientation. Away from the optimal orientation the strongest responses are to low curvatures — stimuli to which there are no responses at the optimal orientation. The shorter excitatory RF responds over a broader range of orientations to low curvature stimuli leading to the strong response to oblique lines. As indicated previously (6) under these conditions an equal spatial frequency ES instance approximates one of the linear RF instances. Hence long stimuli will produce bimodal orientation-tuning curves.

An alternative approach is to choose the parameters so that the orientation bandwidth is the



Figure 8 About Here

same for the two components. Equation A6 implies that a match is accomplished by satisfying:

$$\frac{\Delta v^S}{u_o^S} = \frac{\Delta v^L}{u_o^L} \quad (8)$$

and Figure 6c depicts the situation graphically. Figure 7d and 7e shows the curvature response as a function of line width. Because of the difference in component spatial frequency tuning the shape of the curve is variable. Figure 7f shows the response for curves of different orientations with respect to the RF. Although the response curves vary somewhat in shape, they decline in magnitude with increasing orientation deviation and response does not recover strongly at low curvatures.

Instead of matching the components in spatial frequency (u_o and BW_u) or in orientation (BW_θ) it is possible to effect a tradeoff between the two. The problem is how to choose the components so as to maximize the invariance of the curvature response with changes of curve orientation and width. Figure 7g,h,i show the results of an intermediate choice. Appendix 1 gives an example of a functional that one could minimize in choosing a tradeoff. However this functional is not unique and different functionals will lead to somewhat different parameter choices.

Odd-symmetric ES examples are selective for the sign of curvature as well as its magnitude (Dobbins et al., 1987; 1989). Figure 8 shows the response of an ES instance in which both components are odd-symmetric, but in which the parameter values otherwise equal those in Figure 7g,h,i. Essentially only the magnitude of the curvature response varies with curve width (Figure 8a), while the response recovery to low curvatures at oblique orientations is not that strong (Figure 8b). To the opposite sign of curvature there is almost no response at the optimal orientation, while for low curvature diagonals the response approaches the level of response shown

in Figure 8b. Overall the response is the least variable observed.

The Uncertainty Principle imposes a limitation on the ability to match filter characteristics: the length difference which gives rise to curvature-selectivity implies a difference in Δx for the two components. It is certain that for RFs as in Figure 6a (with matched u_0 and BW_u) that the orientation tuning of the small RF cannot be increased to match the large one. With linear components the position, orientation, and curvature tuning cannot be decoupled.

In passing we note that a corresponding difference in the length-response of simple endstopped cells would be expected depending on whether or not the components are matched in spatial frequency. The implication is that in one case asymptotic end inhibition is independent of the width or spatial frequency of the stimulus used and in the other case it is not. An experiment to determine whether end-inhibition varies with bar width is also of interest with respect to the issue of whether endstopped and nonendstopped cells exist along a continuum (Schiller et al., 1976) or divide into distinct classes (Kato et al., 1978). That question has hinged on the distribution of asymptotic end-inhibition observed.

Summary The linearity of the components of the ES model implies that the Uncertainty Principle establishes a limit on the joint space-spatial frequency behaviour of the model. Constructing endstopped neurons from simple cells with different parameter choices leads to different tradeoffs in the stability of curvature-tuning.

4.2.3 Varying Gain and Size

In this section we consider how variation in the relative gains and sizes of the components affects curvature response. This is necessarily preliminary to any effort aimed at designing a family of mechanisms sensitive to different ranges of curvature.

Since endstopping is usually defined in terms of length-tuning, a brief description of the effect of varying gain and component size on length response is given. Increasing the gain of the small

RF neuron has the following effects: response magnitude increases, optimal length increases (toward the short RF length), and the asymptotic response to long bars increases (decreased end inhibition). Similarly, increasing the length of the small RF shifts the peak response to greater length.



Figure 9 About Here

Figure 9 shows the curvature response of the model as the size and gain of the small component is varied. The left column shows the responses in terms of radius of curvature and the right column plots the log of the response against the log of the radius. The latter has the advantage that the curvature and radius of curvature are now related by a left-right reversal. Looking down the left column in Figure 9 it can be seen that the peak radius increases as the small component increases in size. Similarly, increasing the gain of the small component increases the peak radius as well as response magnitude. However, the change in peak location is smaller than is obtained in length-tuning measurements. Hence, over a limited range, increasing the size or gain of the small RF has a similar effect.

The right column of Figure 9 highlights the differences. Changing the small component gain changes the rate of rolloff on the low curvature side whereas changing size changes the response rolloff rate at high curvatures. The effect of changing relative gain is much more dramatic than changing relative size and principally determines the breadth of curvature-tuning. For example, if one increases the gain of the small RF, the curvature response width at half height increases and the end inhibition as measured by length-tuning decreases. All but one of the examples in Figure 9 have half height curvature bandwidths of three to five octaves.

We have also employed a smooth, saturating nonlinearity with variable gain and spontaneous activity in place of simple half wave rectification (Dobbins et al., 1990). If that function em-

ployed, small values of the spontaneous activity parameter can narrow the curvature bandwidth by reducing the response at very high curvatures. On the other hand, saturation biasing can significantly broaden the curvature bandwidth. Response is most stable when saturation is matched for contrast in the two components. Similarly, for cells which exhibit contrast gain control the minimum change in the shape of endstopped curvature-response occurs when the change in gain co-varies in the excitatory and inhibitory components of the endstopped neuron. Mismatch in contrast gain adaptation would produce changes in curvature selectivity as shown for changing gain in Figure 9. Changes in responsivity or noise that are common to both small and large components do not adversely affect ES model performance, but uncorrelated changes will cause perturbation in curvature-tuning curves.

The examples shown in Figure 9 have components matched in spatial frequency response and the results were for curves of only one width. If the components are not matched in spatial frequency response then the way in which curvature-response varies with relative gain depends on the curve width.

Summary Both the relative gain and the relative size of the endstopped model components affect the shape of the curvature response. Relative gain is more significant and determines the curvature bandwidth whereas the relative size is more important in determining the peak of the response curve.

4.3 Curvature as Orientation Change

Consider a curve described locally as curvature parametrized by arc length ($\kappa = \kappa(s)$). The orientation range ($\Delta\theta$) spanned by a section of arc is the integral of the curvature with respect to the arc. The distinction in the curvature preferences between endstopped and endfree cells can be described in terms of orientation range. For arcs at least as long as the maximal RF dimension endstopped neurons prefer large orientation ranges (e.g. $\Delta\theta > 45^\circ$) and nonendstopped cells

small ones. The orientation range distinction is scale-independent and hence more general than the low/high curvature distinction.



Figure 10 About Here

Figure 10 shows the results of experiments with an endstopped simple model cell in which both curvature and arc length are systematically varied. Figure 10a shows the response surface in radius of curvature–arc length space. Figure 10b plots level sets (iso-response contours) in the same coordinates. For short arcs, the response gradient is in the direction of changing arc length, while for long arcs the gradient is in the curvature direction.

5. Discussion

5.1 The Range of Curvature Selective Cells

Wilson (1985) has argued that the effective orientation range employed in human curvature discrimination is primarily determined by the tuning of the underlying orientation selective mechanisms. If endstopped cells are concerned with representing the local bending of arcs, we now ask how many (with distinctly different stimulus selectivity) are required.

There are two factors to consider: (i). the range of curvature that is represented by single RFs, and (ii). the curvature resolution (inverse of bandwidth) attained by the RFs. There are two constraints affecting the range of representation by RFs. The first is a maximum curvature constraint which is ultimately determined by optical resolution and photoreceptor density. Then the maximum curvature to which the system is sensitive is determined by the smallest orientation selective units. In the primate, a substantial fraction of foveal cortical units is less than 10 arc

minutes in length and the smallest are less than 5 arc minutes (Parker & Hawken, 1988). In cat striate cortex, the neurons with the shortest RFs are almost invariably endstopped (Kato et al., 1978), and since there is substantial or total overlap between RF center and endzones (Orban et al., 1979b), the excitatory RF length of endstopped cells is systematically underestimated (Dobbins et al., 1989). Experiments with the ES model show that the peak of the length response curve may be as little as 50-60 percent of the length of the short excitatory RF. In addition, we have conducted other simulations to determine the relationship between the peaks of the length and radius of curvature tuning curves. The results indicate that the peak radius of curvature varies between the length-tuning peak (typically 0.6 – 0.75 of the small RF length) to an upper value equal to the small RF length. Therefore, if one takes a value of 5 arc minutes as the smallest measured length, assuming the measured length is 0.6 – 0.75 the length of the discharge region, and using the just-described relation between length and radius of curvature one obtains an estimate of peak radius of 5 – 9 arc minutes for the smallest units.

Parker and Hawken (1988) estimated RF lengths indirectly from spatial frequency and orientation bandwidth measurements in monkey. Two thirds of the RFs are within a factor of five in length although the full range was about a decade (less than 5 to greater than 40 arc minutes). The data of Kato et al. (1978) and of Gilbert (1977) also indicate that the range of RF lengths at a particular eccentricity in cat Area 17 is no more than a decade. However since these data in fact represent a range of eccentricities of a few central degrees, and there is greater variability in length measurements than for many other RF parameters, it is possible that the range is rather less.

There is also a kind of minimum curvature constraint. Although it is possible to choose RF parameters to obtain endstopped RFs responsive to low orientation ranges (but still curvature bandpass), a proportionally denser sampling of orientation is implied, making it more economical to use endfree (curvature lowpass) neurons to represent low orientation ranges.

Curvature resolution is not primarily determined by the curvature bandwidth to a particular idealized stimulus but rather by the variation in bandwidth over the class of perturbed curves. Proceeding naively, by combining adjacent elements in a series of four octave-spaced RFs one obtains three endstopped cells. The four cells span a factor of eight in length. The endstopped cells composed in this way have constant log bandwidth and cover a factor of four in their curvature tuning peaks. First note that the model predicts half height curvature bandwidths of 3-5 octaves. This breadth suggests that octave RF spacing represents significant oversampling of curvature. In other words, such sampling does not partition curvature into orthogonal classes. Second, the shortest and longest RFs (at any retinal eccentricity) are very unlikely to be tuned to the same spatial frequency, implying that, for a given eccentricity and preferred spatial frequency there do not exist endstopped cells tuned for curvatures differing by as much as a factor of four. The qualification to be made is that we are referring to cells that are strongly endstopped; cells with weak end inhibition extend the range of peak curvatures but have such broad curvature-tuning that it is unlikely that they play a significant role.

To summarize, examination of the range of RFs available for synthesizing endstopped neurons suggests that the range of peak curvatures represented by endstopped cells is not much greater than a factor of two (for a given preferred spatial frequency). Within this range there may be cells with a range of curvature-tuning characteristics. However, these would not appear to fall neatly into orthogonal curvature classes.

5.2 Linear and Quasi-Linear Endstopped Models

The endstopped simple model exhibits the right qualitative length, orientation and curvature response. As illustrated in Figure 3b plotting λ_y (fraction of the total RF that is excitatory) against end inhibition leads to a distribution very much like the empirical one found in Kato et al. (1978). If the spatial extent of the afferent pool that may contribute to RF synthesis is

fixed, then the linear model predicts that endstopped cells have broader orientation tuning than their nonendstopped counterparts, and that the degree of endstopping is intimately linked to the length of the excitatory RF. Both these results are in accord with the experimental findings of Kato et al. (1978).

Although the linear and quasi-linear ES models exhibit the right qualitative behaviour they can be discounted as plausible endstopped neuron models on empirical grounds: the phase-specific nature of the endzones. Tanaka et al. (1987) have examined endstopped cells in cat Area 19, finding that the endzones are tuned to the same spatial frequency as the RF center but have broader bandwidth. If the endzones were synthesized from one simple cell this would imply that the endzones are narrower than the center. This is incompatible with the evidence of Orban et al. (1979b) on endzone width in cat striate endstopped neurons. In addition, Tanaka et al. (1987) found that the endzones of the Area 19 cells were phase-insensitive implying a complex-like character as first suggested by Hubel and Wiesel (1965). In the accompanying paper we explore models of endstopped neurons with complex components. In addition to nonlinear components one can imagine nonlinear combination of the components. For example one could treat the end inhibition as divisive rather than subtractive. This AND-NOT combination (Koch et al., 1983) can be interpreted as an assertion of the form: the contour can be approximated by a line segment of length L but not of length $2L$. Employing divisive rather than subtractive inhibition changes the shape of the length and curvature response curves somewhat but not the basic character of the model.

Both the linear and quasi-linear ES models respond to chevrons, crosses and so forth in addition to smooth curves. The RF symmetries imply that these models respond to two differently oriented chevrons and hence smooth curves at orthogonal orientations and with distinct curvatures. This undesirable artifact is partly a consequence of the position or phase-specificity of end inhibition. If, instead, inhibition is position-independent in the end zones as the evidence

suggests then the response to curves and chevrons rotated 90 degrees with respect to the center is abolished.

All of the versions of the endstopped model (in this and the accompanying paper) respond to curves, none responds exclusively to curves, and none provides a precise estimate of curvature. Therefore the endstopped simple model is not a 'curvature detector'. Nevertheless, as shall be discussed in the following paper the curvature-selective response of endstopped neurons can be exploited as a first stage in curve inference. The degree of success of the linear and quasi-linear models is attributable, we believe, to more realistic non-linear models resembling the linear model in some range of operation. This can arise from nonlinear components or more nonlinear interactions between them or both (Iverson and Zucker, 1990; Dobbins *et al.* 1990).

6. References

- Bolz, J., & Gilbert, C.D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature* 320: 362-365.
- Bracewell, R. (1965) The Fourier transform and its applications. McGraw-Hill, New York.
- Daugman, J. (1980) Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res.* 20: 847-856.
- Daugman, J. (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional cortical filters. *J. Opt. Soc. Am. A* 2, 1160-1169.
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1987) Endstopping in the visual cortex as a substrate for calculating curvature. *Nature* 329, 438-441.
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1989) Endstopping and curvature. *Vision Res.* 29(10): 1371-1387.
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1990) Endstopped Visual Cortical Neurons II. Experiments with a Complex, Nonlinear Model. (companion paper)
- Gabor, D. (1946) Theory of communication. *Proc. Inst. Elec. Eng.* 93: 429-441.
- Gilbert, C.D. (1977) Laminar differences in receptive field properties of cells in cat primary visual cortex. *J. Physiol. (Lond)* 268: 391-421.
- Heggelund, P. (1986) Quantitative studies of the discharge fields of single cells in cat striate cortex. *J. Physiol. (Lond)* 373: 277-292.
- Hubel, D.H., & Wiesel, T.N. (1965) Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.* 28: 229-89.
- Iverson, L. & Zucker, S.W. (1990) Logical/Linear Operators for Measuring Orientation and Cur-

- vature. Technical Report 90-6, Research Center for Intelligent Machines, McGill University
- Jones, J.P. & Palmer, L.A. (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**: 1187-1211.
- Kato, H., Bishop, P.O., & Orban, G.A. (1978) Hypercomplex and simple/complex cell classifications in cat striate cortex. *J. Neurophysiol.* **41**: 1071-1095.
- Koch, C., Poggio, T., & Torre, V. (1983) Nonlinear interaction in a dendritic tree: localization, timing and role in information processing. *Proc. Nat. Acad. Sci. USA* **80**: 2799-2802.
- Kulikowski, J.J., Marcelja, S., & Bishop, P.O. (1982) Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biol. Cybern.* **43**: 187-198.
- Marcelja, S. (1980) Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.* **70**: 1297-1300.
- McGuire, B.A., Hornung, J-P., Gilbert, C.D., & Wiesel, T.N. (1984) Patterns of synaptic input to layer 4 of cat striate cortex. *J. Neurosci.* **4**: 3021-33.
- Orban, G.A., Kato, H., Bishop, P.O. (1979a) End-zone region in receptive fields of hypercomplex and other striate neurons in the cat. *J. Neurophysiol.* **42**: 818-32.
- Orban, G.A., Kato, H., Bishop, P.O. (1979b) Dimensions and properties of end-zone inhibitory areas in receptive fields of hypercomplex cells in cat striate cortex. *J. Neurophysiol.* **42**: 833-49.
- Parker, A.J. & Hawken, M.J. (1988) Two-dimensional spatial structure of receptive fields in monkey striate cortex. *J. Opt. Soc. Am. A* **5**: 598-605.
- Schiller, P.H., Finlay, B.L., & Volman, S.F. (1976) Quantitative studies of single-cell properties in monkey striate cortex I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.*

39, 1288-1319, 1976.

Tanaka, K., Ohzawa, I., Ramoa, A.S., & Freeman, R.D. (1987) Receptive field properties of cells in area 19 of the cat. *Exp. Brain Res.* 65: 549-558.

Watt, R.J. & Andrews, D.P. (1982) Contour curvature analysis: Hyperacuties in the discrimination of detailed shape. *Vision Res.* 22: 449-460.

Wilson, H.R. (1985) Discrimination of contour curvature: data and theory. *J. Opt. Soc. Am. A* 2: 1191-1199.

7. Appendix 1.

The formulation of Gabor (1946) was introduced to vision by Marcelja (1980) and Daugman (1980) and is used throughout the paper in representing linear RF components because of its simplicity and convenience in both spatial and spatial frequency domains. Recall that Gabor functions are the product of Gaussians with sinusoids. Therefore in general a 2D Gabor RF can be described as:

$$k(x, y) = \frac{1}{2\pi\rho\sigma_{x1}\sigma_{x2}} e^{-\frac{1}{2}\left[\left(\frac{x_1-x_{o1}}{\sigma_{x1}}\right)^2 + \rho x_1 x_2 + \left(\frac{x_2-x_{o2}}{\sigma_{x2}}\right)^2\right]} e^{j[(u_o x_1 + \alpha) + (v_o x_2 + \beta)]} \quad (A1)$$

where (x_{o1}, x_{o2}) gives the center of the Gaussian envelope, σ_{x1}^2 and σ_{x2}^2 the variances and ρ the covariance of the Gaussian. The complex modulation frequencies are given by u_o and v_o and the phases by α and β . Employing a local coordinate system (x, y) centered on the RF and oriented so that the modulation is in the x direction, we consider the following Gabor functions:

$$k_e(x, y; \sigma_x, \sigma_y; u_o) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]} \cos(u_o x) \quad (A2a)$$

$$k_o(x, y; \sigma_x, \sigma_y; u_o) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]} \sin(u_o x) \quad (A2b)$$

which have purely even and odd phase modulation although the evidence indicates that simple cortical RFs exist with all intermediate phases as well (Heggelund, 1986). We assume that $\sigma_y \geq \sigma_x$ and that the modulation frequencies are low. In particular u_o is low enough that the number of lobes in the RF is less than or equal to five ($4\sigma_x \leq 2.5\lambda_x$). The spectrum of the Fourier transform is also shown and is a pair of Gaussians shifted by the modulation frequency u_o , with aspect ratios that are the inverse of those in the spatial domain.

Allowing cosine modulation in the y -direction one obtains something like endstopped RFs:

$$k_{ee}(x, y; \sigma_x, \sigma_y; u_o, v_o) = k_e(x, y) \cos(v_o y) \quad (A3a)$$

$$k_{oe}(x, y; \sigma_x, \sigma_y; u_o, v_o) = k_o(x, y) \cos(v_o y) \quad (A3b)$$

Note that a trigonometric identity allows one to treat each of these doubly-modulated RFs as sums of RFs with 1D modulation. Since there is no evidence of more than three zones along the RF length the spatial wavelength of modulation ($\lambda_y = 2\pi/v_o$) is constrained to be less than or equal to the RF length ($\lambda_y \leq 4\sigma_y$). In Figure 2b v_o is increased from 0 and the spatial envelope takes on the shape shown. The spectrum of the Fourier transform is represented schematically on the right. It is worth noting that spatial frequency has usually had a one dimensional interpretation in neurophysiology and psychophysics. For rectilinear stimuli, this is satisfactory, but disks and rings have a radial organization of spatial frequency (with Bessel kernel; Bracewell, 1965). In referring to spatial frequency bandwidth (BW_u) we shall mean the bandwidth in the direction of the modulation vector (u_o).

The Fourier transform of a long thin vertical bar has a spectrum concentrated along the u -axis. A form which is rotated in space has its Fourier transform rotated through the same angle. In symbols:

$$f(x, y) = f(r, \theta) \Leftrightarrow F(u, v) = F(\rho, \gamma). \quad (A4)$$

where (r, θ) and (ρ, γ) are polar representations in the space and spatial frequency domains, respectively. Rotating the line through α :

$$f(r, \theta + \alpha) \Leftrightarrow F(\rho, \gamma + \alpha) \quad (A5)$$

Therefore the Fourier transform of a rotated line is the rotated transform of the original line. A

vertical sinusoidal grating transforms to a pair of impulses on the u axis and rotating the grating sweeps out a circle of radius u_0 in the u - v plane. The preferred orientation of an RF is normal to the direction of the modulation vector, hence the location of the envelopes in the spatial frequency plane in Figure 2a and 2c provides an indication of preferred orientation in response to lines and gratings. The orientation bandwidth of a simple cell viewed in the frequency domain can be treated as:

$$BW_\theta = \tan^{-1}(\Delta v/u_0) \quad (A6)$$

where Δv is the variance of the RF in v (Kulikowski *et al*, 1982), or, approximated graphically by the arcs from the origin to the RF envelope as shown in Figures 6a and 6c. This relation can be expressed equally well in terms of the corresponding spatial domain variables ($1/\lambda_x$ and $1/\Delta y$).

Having obtained a measure of orientation bandwidth we derive a measure of the joint spatial frequency-orientation match of two simple RFs of different size. The degree of match in spatial frequency of the simple components can be represented by the overlap of the two Gaussian kernels in the frequency domain. Therefore the component matches vary with the ratio:

$$\frac{|\Delta v^S/u_0^S - \Delta v^L/u_0^L|}{\iint K^S(u, v) \cdot K^L(u, v) du dv} \quad (A7)$$

where the integration is taken over the plane (\mathbb{C}^2 is reduced to \mathbb{R}^2 for equal phase RFs) and we desire small values as the component parameters u_{0S} , u_{0L} , Δu_S , and Δu_L vary.

8. Figure Legends

Figure 1. a. The endstopped model is constructed from the convergence of excitation from a short RF cell and inhibition from a long RF cell.

Figure 2. a. Even and odd (k_e, k_o) Gabor representations of simple RFs in space (left) and their Fourier spectrum (right). The ellipses have Gaussian weighting in both domains. b. The transition between the RFs in a. and c. as v_o is increased from 0 in both domains. The arcs bounding the envelopes in the frequency domain provide a graphical indication of the orientation bandwidth in response to very long lines. Similarly, the migration of the envelopes away from the u axis is related to the change in asymptotic end inhibition measured with long lines (or spatial DC response). c. Simulating endstopped RFs with additional cosine modulation in the y -direction (k_{ee}, k_{oe}). Note that in response to long bars, the preferred orientation has shifted and the orientation bandwidth has increased.

Figure 3. a. Length-response curves for the k_{ee} Gabor RF for different values of v_o (0, 0.3, 0.6, 0.9). (RF parameters: Size: 21 pixels, Aspect Ratio: $\sigma_y/\sigma_x = 2$, $u_o = 0.15$) As the inhibitory end zones (modulation frequency) increase, the asymptotic inhibition goes from 0 to 100 percent. b. The asymptotic end inhibition as a function of the relative lengths of the excitatory and inhibitory zones. The abscissa is the wavelength of the modulating sinusoid (λ) normalized by the RF length (S).

Figure 4. a. The orientation response for k_{ee} instances with the same parameter values as the previous figure. The response curve bifurcates as v_o increases. b. Orientation response for short (solid curve) and long bars (dotted curve) of width 3 pixels for the instance with $v_o = 0.06$. The short bars equal the excitatory center in length and the long bars the total RF length. Although unimodal, the response to short bars has broader bandwidth than the for the nonendstopped instance in response to long bars.

Figure 5. a. The response to appropriately oriented semi-circular arcs for the same set of

parameter values as previously. For low values of v_o the curves appear nonendstopped but become increasingly endstopped as v_o increases. b. The data are replotted on log-log axes. The curves differ in their characteristics both on the low and high curvature side. Plotted this way curvature and radius of curvature are merely related by a sign change.

Figure 6. The RFs and Fourier spectra of different pairs of vertically-oriented simple Gabor RFs that differ in length (after Daugman, 1985). a. Same bandwidth in u (BW_u) and center frequency (u_o) b. Equal envelope aspect ratio (AR) and u_o . c. Equal BW_u and orientation bandwidth (BW_θ). d. Equal center frequency u_o .

Figure 7. The curvature response as a function of curve width and orientation for three ES instances. The left and middle columns show response to curves of different width plotted on linear and log axes, respectively. The line widths are 1, 3, 5, and 7 pixels (solid, dotted, dashed, long-dashed). The right column shows the response to curves of width 3 pixels at orientations from optimal to 25 degrees from optimal in steps of 5 degrees. In all cases the radius of curvature varies from 0 to 200 in steps of 10. the top row (a., b., c.) corresponds to Figure 6a (parameters: Size-S/L: 35/61, Aspect Ratio (AR) S/L:1.15/2.0, center frequency u_o -S/L:0.05/0.05, and gains c^S/c^L :1.8/1.0) designed so that the components are matched in center spatial frequency (u_o) and bandwidth (BW_u). a. The response as a function of curve width. b. On log-log axes the curves appear as vertically translated copies. c. There is strong response to high curvatures at the vertical orientation and low curvatures at oblique orientations. The middle row (d., e., f.) corresponds to Figure 6c (parameters: Size-S/L: 35/61, AR-S/L:1.15/2.0, u_o -S/L:0.08/0.046, c^S/c^L :1.9/1.0) designed so that the components are matched in spatial frequency (BW_u) and orientation (BW_θ) bandwidth. d., e. The response as a function of curve width is variable and there is no response to the widest curve. f. The response narrows and decays as the curve orientation moves away from vertical. There is no response to low curvature at oblique orientations. The bottom row (g., h., i.) represents one choice of tradeoffs between the

matching in spatial frequency and in orientation (parameters: Size-S/L: 35/61, AR-S/L:1.15/2.0, u_o -S/L:0.06/0.046, c^S/c^L :1.85/1.0). g., h. Comparing with the panels above it is clear that the form of the response is less sensitive to curve width than in d and e but more so than in a and b. i. The reverse is true for sensitivity to curve orientation.

Figure 8. The curvature response as a function of curve width and orientation for an odd-symmetric ES instance. The left and middle graphs show response to curves of different width plotted on linear and log axes, respectively. The line widths are 1, 3, 5, and 7 pixels (solid, dotted, dashed, long-dashed). The right graph shows the response to curves of width 3 pixels at orientations from optimal to 25 degrees from optimal in steps of 5 degrees. In all cases the radius of curvature varies from 0 to 200 in steps of 10.

Figure 9. The curvature response of three ES instances as the gain of the small component is changed. The right column re-expresses the data of the left column on log-log (base 2) plots. In each case 51 curves of radius from 4 to 204 pixels and width 3 pixels were used. In a., b., and c. the small components are of length 21, 31 and 41 pixels, respectively, with AR 1.0, 1.4, and 1.6. In each case, for both small and large components $u_o = 0.05$. The large component is always the same – 61 pixels long with AR = 2.0, and $c^L = 1.0$. a. For the 21 pixel RF c^S takes values 2.5, 3.0 and 3.7 (solid, dotted, dashed) for the three curves shown. b. a. For the 31 pixel RF c^S takes values 1.8, 2.1, and 2.4 for the three curves shown. c. For the 41 pixel RF c^S takes values 1.3, 1.6 and 2.0 for the three curves shown. In each case the relative gain determines the rolloff on the right (low curvature) side. The relative size determines the rolloff on the left (high curvature) side and the peak position. The locus of the peak radius is approximately equal to the length of the small RF.

Figure 10. The response of an endstopped simple instance to curves in which arc length and radius of curvature are systematically varied. The parameters are as in Figure 11a,b,c. a. Three dimensional response plot. The termination of the surface from the left foreground shows the

curvature response curve while a view from the right shows the (reversed) length-tuning curve.

b. Level sets of the response surface. For short arc lengths there is low curvature selectivity and high length selectivity, while for long arcs the reverse is true. This is evident by examining the direction of the response gradient (normal to the level set contours).

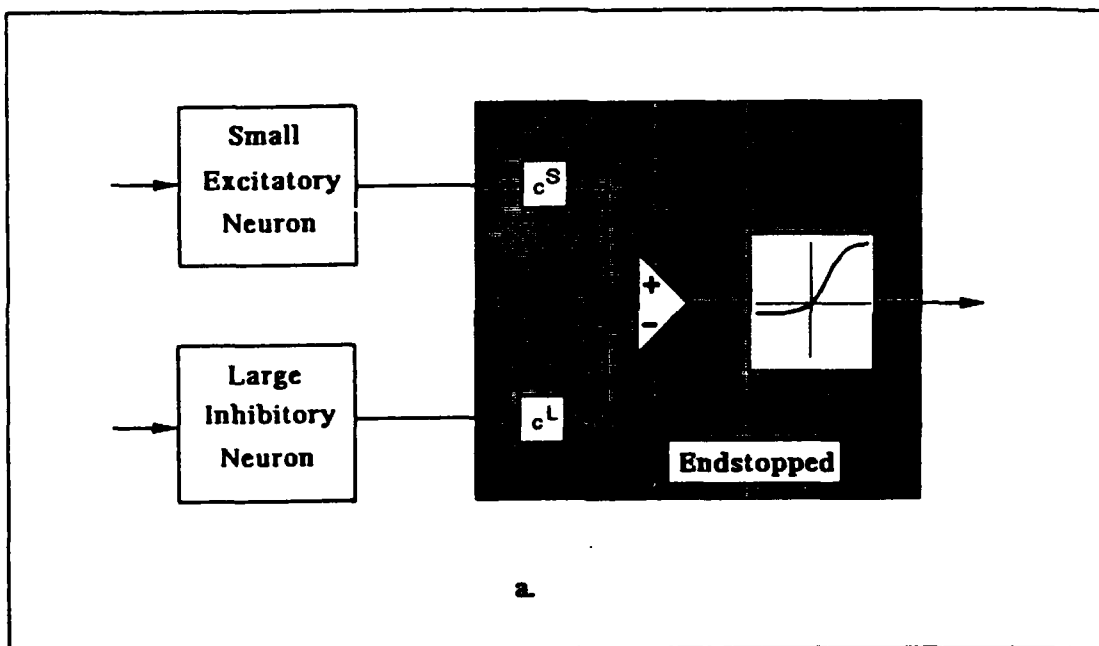


Figure 1

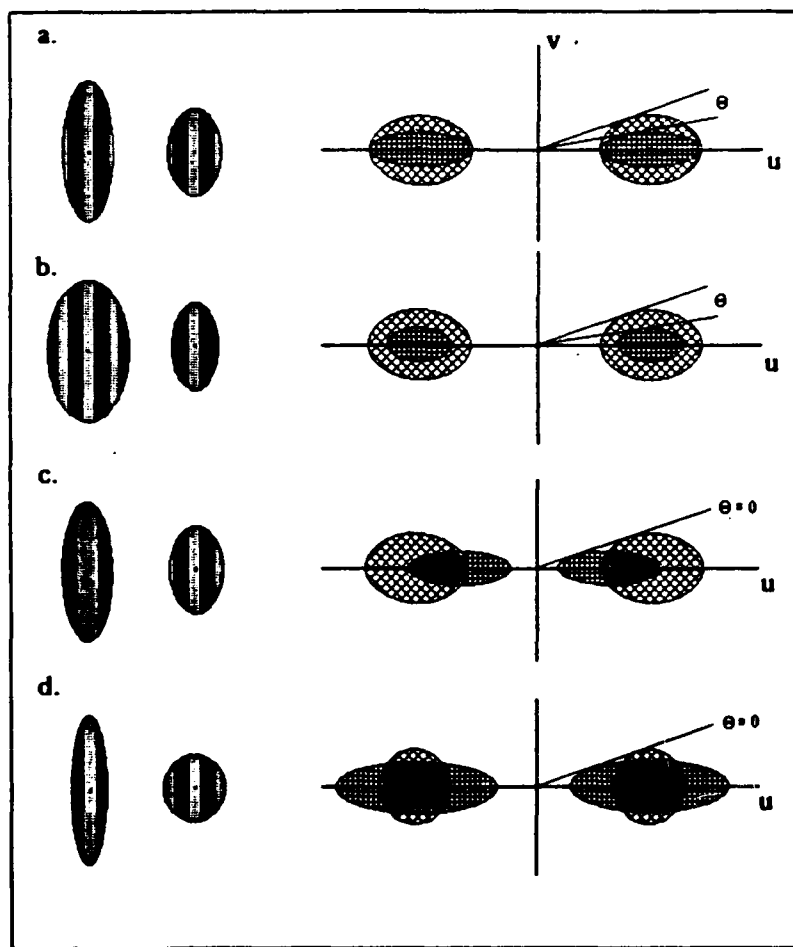
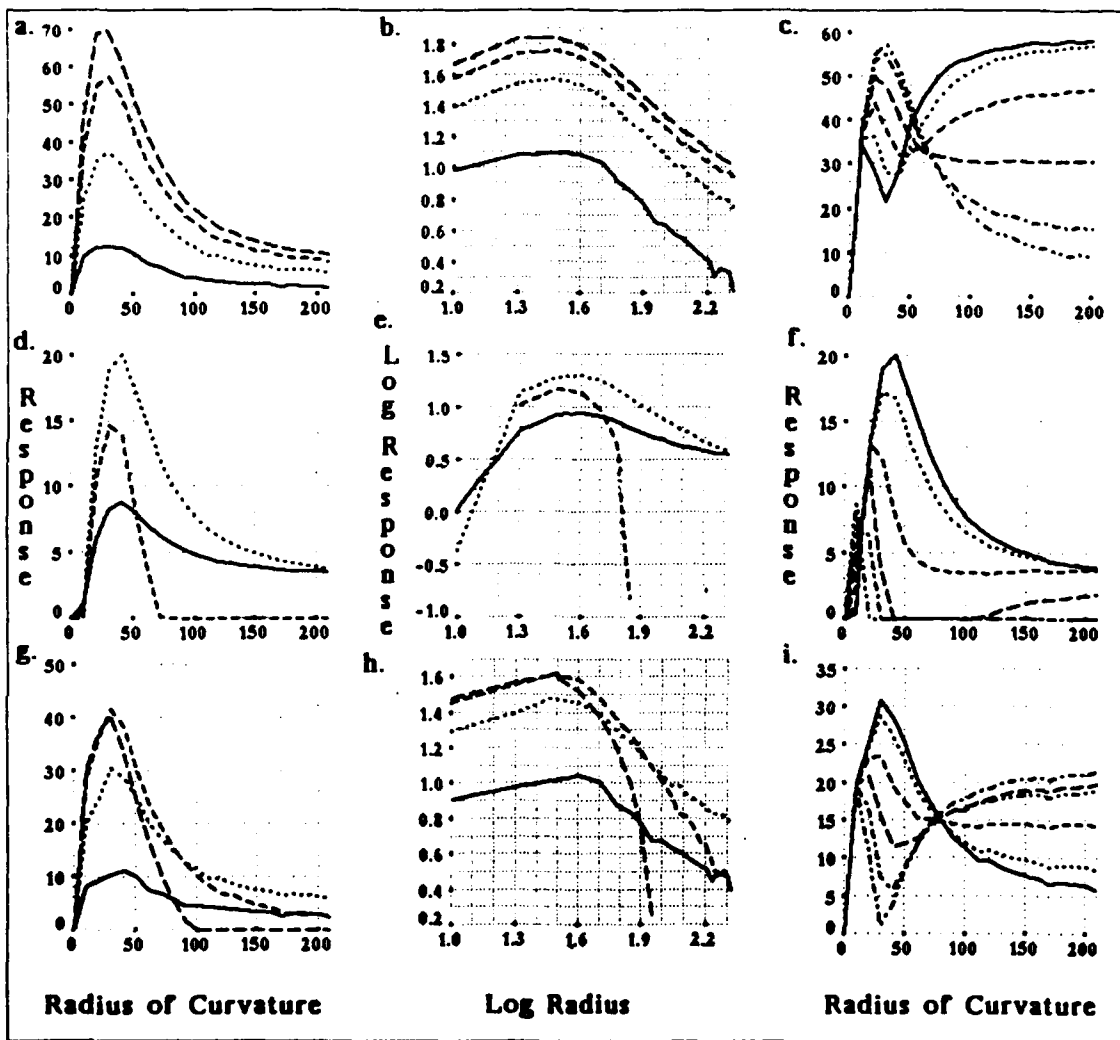
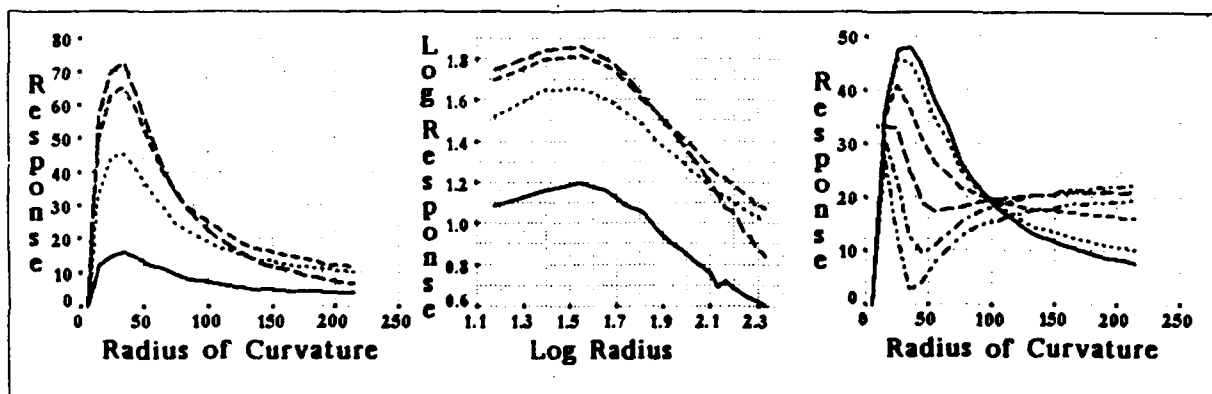
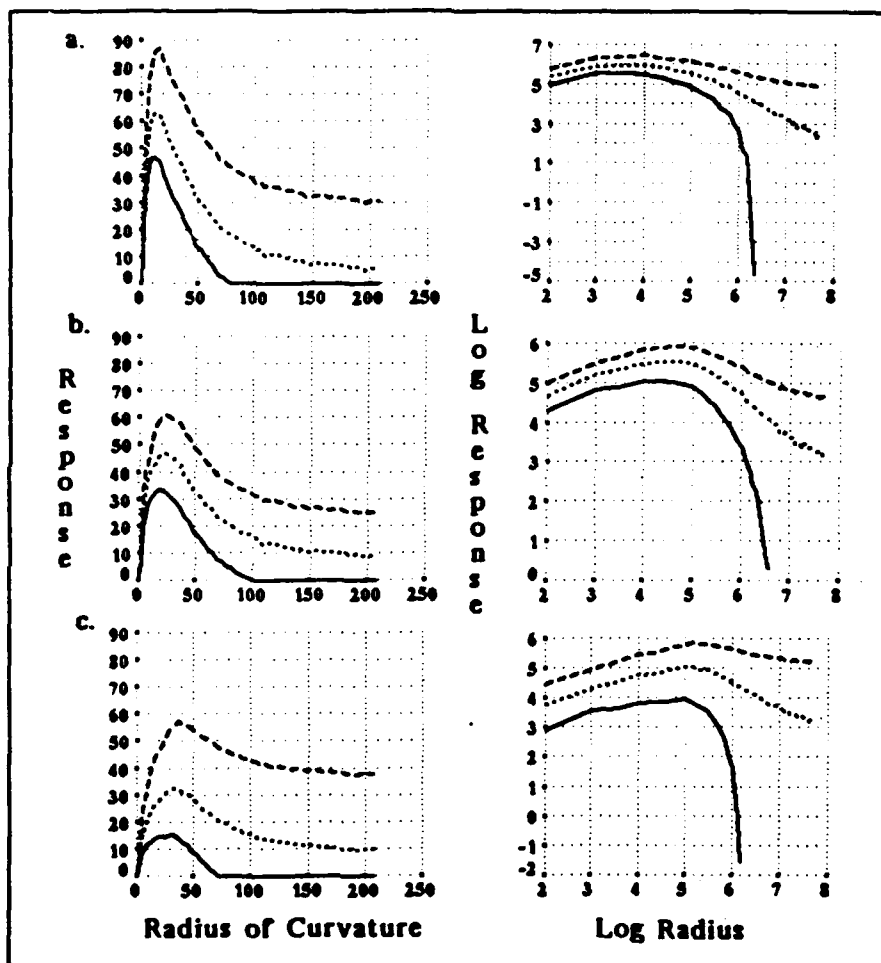
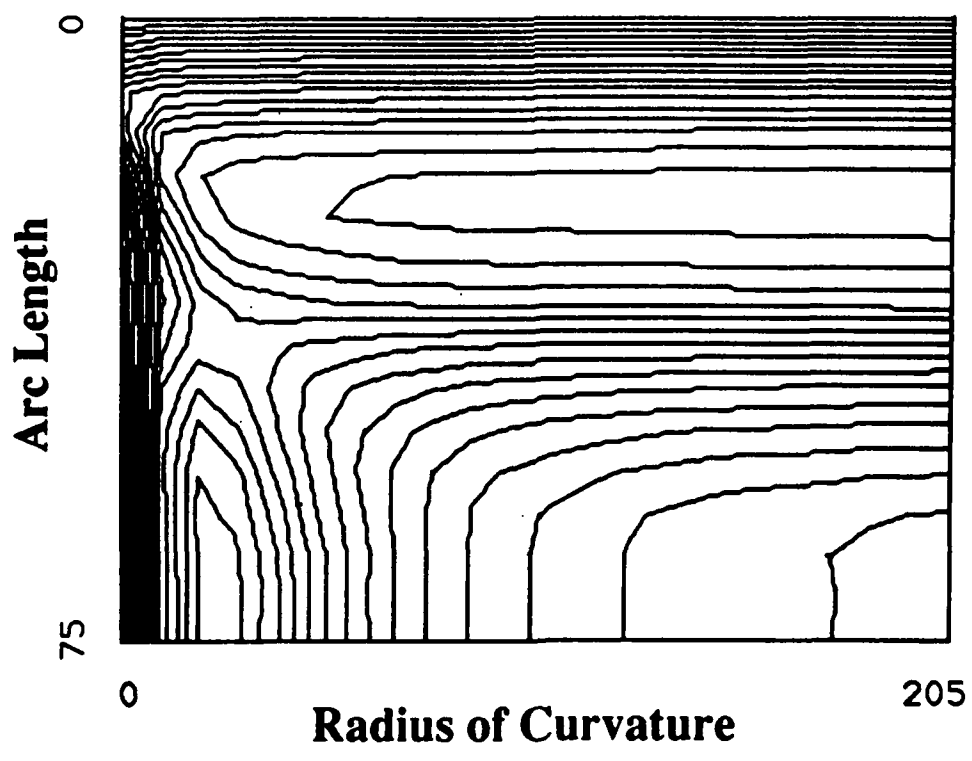
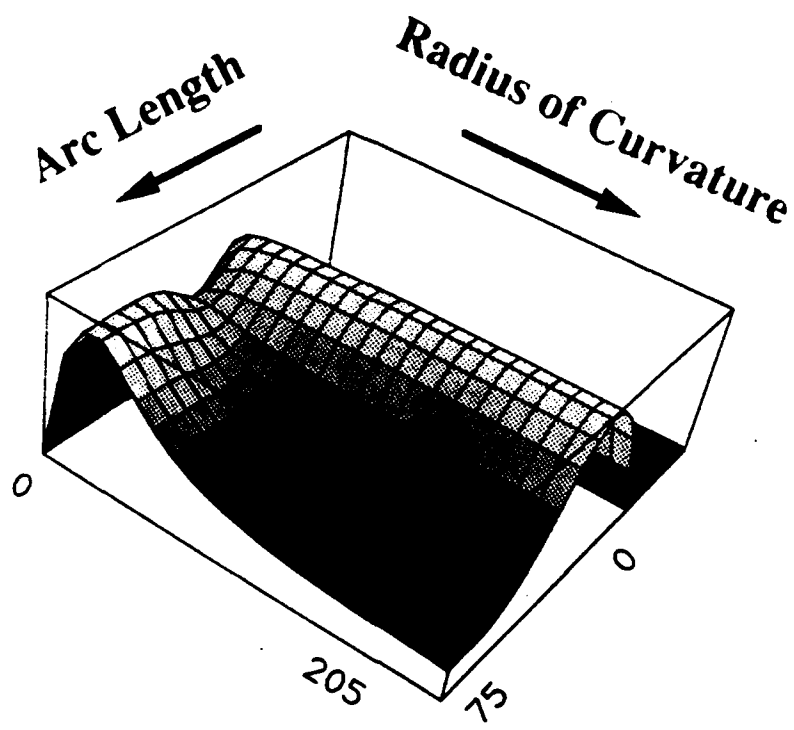


Figure 6









Appendix IV

Endstopped Visual Cortical Neurons II: Experiments with a Complex, Nonlinear Model.

Endstopped Visual Cortical Neurons II. Experiments with a Complex, Nonlinear Model.

Allan Dobbins Steven W. Zucker * Max S. Cynader †

Computer Vision and Robotics Laboratory
McGill Research Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada

* Fellow, Canadian Institute for Advanced Research

*

† Dept. of Ophthalmology, University of British Columbia, Vancouver, B C Canada

Running Title: Experiments with a Complex Endstopped Model

27 pages 7 figures 0 tables

Postal Address: 3480 University Street, Montréal, Québec, Canada H3A 2A7

Telephone: (514) 398 6319 Telex: 05 268510 FAX: (514) 398 7348

Network Address: mrcim@larry.mrcim.mcgill.edu

Endstopped Visual Cortical Neurons II. Experiments with a Complex, Nonlinear Model.

Allan Dobbins Steven W. Zucker Max S. Cynader

Abstract

A model of complex endstopped neurons is developed and analysed. The endstopped neuron model combines an excitatory center which may be simple or complex and complex inhibitory end zones. Complex cells are classified as standard or special according to positional, length-summing and orientation-tuning criteria. The simulations reported here suggest there may be two factors contributing to the special complex behaviour: (i) subadditive synaptic interactions among inputs, and (ii) the presence of endstopping. The curvature selectivity of the endstopped complex model is evaluated with curved stimuli that are varied in orientation and width to examine the robustness of response to stimulus perturbation. The results depend strongly on how the parameters of the constituents of the model cell are chosen. In some cases one obtains curvature-selective responses very similar to those obtained with the endstopped simple neuron model. In some other cases the behaviour is more complicated. Finally, we consider the range of curvature spanned by endstopped cells and discuss representational issues in curve description.

Acknowledgements

We wish to thank Lee Iverson, Benjamin Kimia and Peter Whaite for valuable discussion and programming support, and Patrick Cavanagh for helpful comments on the manuscript. Research supported by NSERC grant A4470 to the second author, and by NSERC grant A9939 to the third author, and by AFOSR grant 89-0260.

1. Introduction

Descriptions of visual objects are related to descriptions of their surfaces and bounding curves. The processes underlying these share important features. In neither case can continuity, linearity, smoothness, or freedom from occlusions be blithely assumed. Instead, one needs information from the visual image concerning whether and where these are true. To estimate the structure of curves one must answer: Is it a curve (a continuous line or edge)? Is it straight or curved? Does it curve smoothly, and, if so, how rapidly? It follows that a necessary prerequisite for a theory of curve description is determination of which of these questions can be addressed locally, and which require interactions across distance. The present study maintains a narrower focus, wherein we continue exploring the hypothesis that endstopping is a property related to the estimation of curvature, and evaluate models of endstopped neurons and their success in estimating curvature.

Dreher (1972) established that the discharge region of endstopped cells may be either simple or complex, but a variety of evidence implies that the endzones have a complex character (Hubel and Wiesel, 1965; Orban et al., 1979a, 1979b; Tanaka et al., 1987). Therefore in this paper we extend the endstopped simple model (Dobbins et al., 1987; 1989; 1990) to include complex components. The complex model parameters can be chosen to resemble a cell of either the standard or special complex type. Length, orientation, and positional sensitivities of the complex cell model are computed to establish a baseline of comparison for simple cells and the experimental data. Simulations are then performed in which the curvature-selectivity of the endstopped complex model is evaluated.

2. Methods

Simulations were developed and run on a Symbolics 3670 LISP computer. Most of the experiments reported here examine the response of the model to curved arcs of different radii with varied curve position, orientation or width. A complete description of the simulation methods

Figure 2 About Here

is contained in Dobbins et al. (1990).

3. The Endstopped Model

Figure 1 About Here

The essential character of the endstopped model is that its response arises from the "difference" in response of cells with receptive fields (RFs) of similar position and orientation but different size. The small (S) and large (L) cells can be simple or complex (Figure 1). Notation: superscripts (S , L , ES) refer to particular cells while subscripts (e , o) refer to even or odd RF symmetries. Each convolution result (R^S , R^L ; $R^i = k^i * I$) is passed through a function ($\Phi(\cdot)$) which models the inability of neurons to represent negative values on a low spontaneous firing baseline. Scalar constants c^S and c^L are introduced to balance the responses between S and L . Therefore the response of the endstopped cell is given by:

$$R^{ES} = \Phi(c^S \cdot \Phi(R^S) - c^L \cdot \Phi(R^L)). \quad (1)$$

The activation (or positive part) function Φ will be used to clarify the relationship between certain linear and nonlinear models. Previously, simple half wave rectification was employed for this purpose (Dobbins et al., 1987; 1989).

3.1 The Activation Function

Observe that Φ depends on two parameters Γ and ρ where Γ determines the form of the nonlinearity and ρ controls the contrast sensitivity. Φ contains a saturating nonlinearity to reflect the dynamic range limits of cortical neurons and Γ models spontaneous activity which determines the degree of rectification (or, equivalently, the amount of compression of negative values). We have modelled Φ as:

$$\Phi(R; \rho, \Gamma) = \frac{1 - e^{-R/\rho}}{1 + 1/\Gamma e^{-R/\rho}} \quad (2)$$

where R is the convolution result, and $\Gamma \in (0, 1]$ (see Appendix 1 for further details concerning Φ). Observe in Figure 2, that for small Γ the output has a saturating half wave rectified character while for $\Gamma = 1$ there is no rectification and the range of output is equal for positive and negative R . In addition, the width of the linear range is roughly independent of Γ . Therefore, apart from quadratic distortion at low R and saturation at high R , Φ with small Γ approaches simple half wave rectification ($\phi(\cdot)$):

$$\lim_{\Gamma \rightarrow 0} \Phi(R; \rho, \Gamma) \approx \phi(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Varying the relative values of Γ and ϕ in relation to the input magnitude allows selection of a variety of different input-output properties. In much of what follows it is assumed that each of the model components has length summation, however we shall also consider special complex cells which do not exhibit length summation and which are often endstopped.

3.2 The Receptive Field Components

The endstopped cell discharge region and endzones may be either simple or complex, and this section reviews the models of simple and complex cells that are employed. The parameters of Φ are selected to produce orientation-tuning and length summation behaviour resembling either standard or special complex cells.

3.2.1 Simple Cells

We model the response of simple neurons as spatial convolutions against the image and represent the RFs by Gabor functions (see Appendix 1. Dobbins et al. 1990). This class of model has been shown to provide a good fit to the spatial RFs of simple cells (Jones & Palmer 1987), although other models also could have served as well (Parker & Hawken, 1988).

3.2.2 Complex Cells

In their hierarchical model of the construction of visual cortical cell types, Hubel and Wiesel (1965) postulated that hypercomplex cells are synthesized from complex cells which in turn are composed of simple components. There is a variety of evidence incompatible with this scheme in its simplest form, such as complex cells being driven monosynaptically from the LGN (Hoffmann & Stone, 1971; Singer et al., 1975), and endstopped (hypercomplex) cells appearing in both simple and complex varieties (Dreher, 1972; Kato et al., 1978). Although cortical neurons have significantly narrower spatial frequency bandwidths than LGN cells, many common experimental assays would fail to distinguish complex cells with staggered simple components from cells receiving a suitable array of LGN afferents. Complex cells behave spatially *as if* they receive input from simple cells with slightly shifted positions. Movshon et al., (1978) described double bar experiments in complex cells in which the sign and magnitude of interaction depended on both the spatial and temporal intervals. They argued that spatial frequency response is determined by RF subunits, the interaction of which determine the directional and velocity preference. Heggelund (1981) modeled the synthesis of complex cells from overlapping ON and OFF center LGN afferents contributing both excitation and inhibition. In contrast, Spitzer and Hochstein (1985) treat complex cell behaviour as arising from sums of half wave rectified simple cells. Their model, which consisted of one spatial dimension (across the RF) and time, successfully predicts the temporal response of many simple and complex cells to sinusoidal grating stimuli.

Standard Complex

Beginning with the original description of complex neurons (Hubel & Wiesel, 1962), investigators have described their diversity in orientation tuning and length summation. Our present goal leads us to concentrate on the spatial aspect of response dependency in classical complex cells (those with ON and OFF responses everywhere in the RF), and not to attempt to portray the features accounting for speed and directional preference. Therefore we build complex-like cells from laterally displaced simple cells, rectifying and giving an individual weighting to each. The complex cell response is given by:

$$R^{CX} = \sum_{i=1}^n c^i \cdot \alpha(R^i) \quad (4)$$

where R^i is the response of the i th simple cell and c^i its weight. For simplicity the simple cell components are equally spaced (shifted less than 1/4 the spatial wavelength) and Gaussian weighted with position. In a purely spatial model there is no distinction between responses to stimuli being turned on and off. Notice that for a particular line stimulus, only one of a pair of even symmetric simple (or LGN) components with equal and opposite RFs would respond. Therefore, the spatial response due to a pair of half wave rectified components of the same position but with opposite contrast sign is equivalent to the full wave rectified response of either of the pair. For computational expedience, in the simulations we use only ON center simple cells with full wave rectification.



Figure 3 About Here

Figure 3 shows line weighting functions and orientation tuning for a complex model instance and one of its simple components. The complex cell responds over a broader area, although

this is reduced if the simple components are half wave rather full wave rectified. There is no corresponding effect on orientation tuning. Note that the simple and complex orientation curves differ insignificantly in tuning breadth.

The orientation width of the complex model instance is approximately the same as its simple components (Figure 3c,d). This is not in agreement with the population averages reported in many investigations (see Orban (1984), for a review). Among the possible explanations are: (i) complex cells are not synthesized from simple cells, (ii) the simple cells that compose complex cells have broader tuning than the simple population as a whole, (iii) other components of complex cells broaden the orientation-tuning, (iv) subunit interaction broadens orientation-tuning, (v) the simple cell components have randomly perturbed orientation preferences, broadening the complex cell orientation tuning (suggestion due to C.L. Baker, personal communication), (vi) complex cells exhibit higher firing rates than simple cells, and perhaps operate closer to saturation, (vii) the complex population is less homogeneous than the simple population. In particular, special complex cells have considerably broader orientation tuning than simple or standard complex cells which differ only slightly (Hammond & Pomfrett, 1989). There is a paucity of evidence on most of these points, but the last two hypotheses can be examined with the activation function $\Phi(\cdot)$ in the simulations. By setting the gain and saturation appropriately one can broaden orientation tuning, reduce length summation and examine the consequences for the endstopped model.

Special Complex

We briefly describe the characteristics of special complex neurons (Gilbert, 1977), and illustrate how the complex cell model can produce them. Special complex cells are characterized by lack of length summation and broad orientation tuning. In addition, they often have high spontaneous activity and are frequently endstopped (Gilbert, 1977; Schwark et al., 1986; Weyand et al., 1986; Hammond and Pomfrett (1989)). A complex model cell appears to be of the special type when the gain is chosen so that the activation function saturates. The saturation is actually

subadditivity or occlusion,

$$\Phi(x_1 + x_2) < \Phi(x_1) + \Phi(x_2).$$

and can be biophysically modeled as subadditive synaptic interaction (Rall et al., 1967; Rall, 1970). Schwark et al., (1986) have shown that the supragranular layers are a necessary input to the special but not the standard complex cells of Layer 5 in the cat. Therefore subadditive interactions could occur on the apical dendrite in Layers 2-3, or via the supragranular projection to Layer 5.



Figure 4 About Here

Figure 4 shows length and orientation tuning curves for the complex cell of the previous section when biased to exhibit subadditivity. The length tuning responses saturate at lengths considerably less than the RF length (61 pixels) and the orientation tuning is broadened.

4. Endstopped Model with Simple and Complex Components



Figure 5 About Here

4.1 Simple Center and Complex Endzones

We begin by examining the response of an endstopped cell in which the small cell is simple and the large inhibitory cell is complex. Figure 5 shows the curvature-response of such a cell

as a function of the width of the curve and curve orientation. In Figure 5a it can be seen that there is some variation in response with the width of the curve and in Figure 5b it can be seen that there is some response recovery for low curvatures at oblique orientations. However it is less than half the response to high curvatures at the optimal orientation. The present example has excitatory and inhibitory components matched in center frequency, but achieves robustness with orientation perturbation with increased orientation bandwidth in the inhibitory complex cell. In particular, the envelope aspect ratio and modulation frequency reduce the inhibitory sidebands thus broadening the orientation bandwidth and DC response. The same effect could have been achieved by using the activation function to bias the inhibitory cell to look special complex-like.



Figure 6 About Here

4.2 Complex Center and Complex Endzones

Figure 6 depicts the curvature response of a cell in which both small and large components are complex. Perturbing curve orientation (Figure 6b) leads to somewhat greater oblique low curvature recovery than for the cell in the previous figure. Figure 6d repeats the experiment for a curved edge. There is no response to curves at orientation deviations of greater than 10 degrees while the 0 and 10 degree curves are virtually identical. In Figure 6c a curved line and a curved edge are tested at all positions across the RF. This cell gives strong curvature-dependent responses to both curved edges and lines, although not in a position-independent way. In both cases the maximum response is displaced from the RF center, but the edge response has odd symmetry being maximal when the edge has not intruded as far as the RF center, while the line response is roughly even about the RF center. If one measures the curvature-tuning curve at different RF positions one obtains different results in both the magnitude and shape of the curve.

The model is purely spatial while real neurons have temporal dynamics. We conjecture that the response to a curve swept across the RF in the preferred direction would largely integrate across the variation in the curve line weighting function.

Figure 7 illustrates the length and orientation responses of the same endstopped complex cell. Note that the peak of the length-tuning curve is significantly less than the length of the short excitatory RF (22 compared to 35 pixels). If one evaluates the orientation response with a bar of this length one obtains a curve with width at half height of 6° degrees. In contrast, evaluated with a bar of length 35 pixels the width at half height is 50 degrees, and for a long bar one obtains the bifurcated curve shown. Thus the broader orientation tuning of endstopped cells than their nonendstopped counterparts (Kato et al., 1978) is a natural consequence of the fact that for endstopped cells the measurements are made at shorter lengths relative to the underlying discharge region (see also Dobbins et al., 1989). Similarly, the endzones reduce the apparent length summation and create a certain amount of position-independent response along the length of the discharge region i.e. the optimal bar can be moved up or down a substantial amount in the RF and still elicit a strong response. This is also a common property among the cells of V4, many of which have both antagonistic surrounds and length-dependent suppression *within* the discharge region (Desimone & Schein, 1987).

We have seen that endstopping can broaden measured orientation tuning and reduce apparent length summation. Since special complex cells are frequently endstopped (Gilbert, 1977; Schwark et al, 1986; Weyand et al., 1986; Hammond and Pomfrett 1989), this raises the question of whether special complex cells really form a distinct category. It may be that endstopping is partly responsible for conferring their apparently distinctive position, length and orientation properties (see section on Endstopping Circuitry also). All of the just-cited studies have noted the high spontaneous activity of special complex cells, so in Figure 7c,d the activation function is introduced (with moderate spontaneous activity, $\Gamma = 0.5$) to the output of the small component.

Figure 7c shows the broadened orientation-tuning that results (width at half height 100 degrees)

Figure 7d shows the curvature response with orientation perturbation. The oblique low curvature recovery is substantial. Saturation was not a factor in these results.

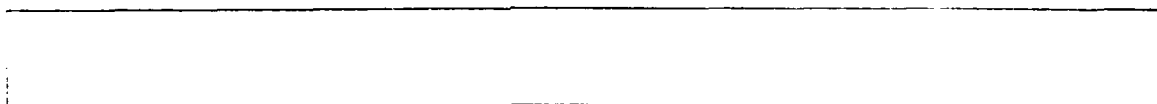


Figure 7 About Here

Employing both complex components and the activation function significantly increases the dimensionality of the parameter space and the range of behaviour of the endstopped model. It also reduces the density of good parameter choices for stable curvature selective response. We found that we could obtain very unusual length, orientation and line weighting functions when combining complex cells to form endstopped cells. For example in one instance the model cell produced a strong bifurcated orientation response (usually diagnostic of bandpass curvature response) but gave almost no response to curved stimuli. We conclude that endstopped complex cells can exhibit stable curvature response but there exists no necessity that they do so.

These explorations imply that endstopped complex cells can produce stable curvature-selective response, however because of the the extraordinarily broad range of characteristics that complex cells can exhibit there are many more ways of failing to achieve stable curvature response than to achieve it. The presence of endstopping broadens measured orientation tuning, causes some longitudinal positional-independence and causes the optimal length to be significantly less than the length of the discharge region.

5. Discussion

5.1 Curve Representation in Visual Cortex

Since Hubel and Wiesel's first explorations in visual cortex it has seemed likely that cells with

oriented RFs are involved in the description of visual contours. In this section we recapitulate two closely related perspectives on local curve representation at the level of striate cortex.

We begin by supposing that a contour is represented locally by a truncated Taylor series approximation. The difference between the two versions depends on whether emphasis is laid on the classical RF (discharge region) or the classical RF plus inhibitory zones. In the first version we suppose that the curve ($\gamma(s)$) is represented locally as:

$$\gamma(s) = \gamma(0) + \gamma'(0)s + R_2(s, 0) \quad (5).$$

where the first two terms are position and orientation estimates at arc position $s = 0$, and the residual R_2 depends on curvature. In this scheme we suppose that the representation of curves is by local line segments and that the purpose of the residual is to determine the largest arc length over which the error in using the linear (tangent-based) approximation is less than some value B . Rearranging and taking norms,

$$|\gamma(s) - \gamma(0) - \gamma'(0)s| < B \quad (6).$$

that is, we wish to find the largest value of s such that the inequality is satisfied. Put differently, the goal is to determine the tangent differential ($\gamma'(0)s$), which entails finding both the tangent and the maximal arc length. In this view discharge regions of different length represent assertions about tangent differentials for different arc lengths, and the asymmetric inhibitory connections from long to short RFs underlie selection of the longest arcs that fit the data. For slowly curving arcs short RFs are suppressed by long RFs which represent the tangent differential, while for rapidly curving arcs long RFs respond weakly and fail to suppress the short RFs which represent the tangent differential. According to this viewpoint curves are represented at this stage as overlapping sets of line segments of different length.

A corollary of this view concerns the asymmetric inhibition from long to short RFs. Clearly,

if a long RF cell at a given position and orientation is strongly active, the activity of shorter RF cells at that position and orientation is redundant. Therefore representing a curve in terms of the longest segments that satisfy some error bound is a more compact representation (see Leclerc (1989) for a development of minimal descriptive length ideas in vision)

In the second version the inhibitory endzones are considered in the RF description. Endfree neurons are described as before while endstopped neurons represent an additional assertion about curvature. Some endstopped cells respond selectively to the sign of curvature of the stimulus instead of simply to the magnitude of curvature (Dobbins et al., 1987; Versavel et al., 1990) supporting the stronger position on curvature, and the proposal that endstopped cells represent a position-tangent-curvature estimate (Dobbins et al., 1989). Although the evidence to date suggests that the majority of endstopped neurons are not selective for the sign of curvature this should be no more worrisome than the fact that the majority of cells preferring motion are not directional. Nevertheless, if the endzones of endstopped cells exhibit phase-independent inhibition as a variety of evidence suggests (Hubel and Wiesel, 1965; Orban et al., 1979b; Tanaka et al., 1987), then it is difficult to account for curvature sign selective responses in a simple way. Among the possibilities are: (i) not all endstopped neurons have complex-like endzones, or (ii) curvature-sign selectivity arises from a difference in temporal properties in discharge region and endzones. A temporal difference would imply that the degree of curvature sign selectivity would be speed or temporal frequency dependent.

The endstopped simple model is closely related to the fuzzy derivative notion of Koenderink and van Doorn (1987) in which the RFs are low order derivatives of Gaussians, and the convolution results represent the Taylor series coefficients (*k*-jets). The principal difficulty with this stronger position is that endstopped cells also respond well to short, uncurved arcs. Indeed endstopped cells respond well to both short uncurved arcs and longer, curved ones. Finally although endstopped neurons respond in a curvature-selective way, the response of an individual neuron varies along

many stimulus dimensions, and the response of an endstopped neuron should not be viewed as delivering a number representing curvature. If interested in explicitly computing curvature, one can combine the responses of RFs using the formulas of calculus (Koenderink & van Doorn, 1987; Koenderink & Richards, 1988). However, it is our view that the responses of endstopped and endfree cells are employed in curve inference networks in the cortex in such a way that their orientation and curvature-selectivities are exploited, but without explicitly or precisely computing curvature (Zucker et al., 1989).

Both simple and complex endstopped neurons exhibit curvature-selective responses (Dobbins et al., 1987; Versavel et al., 1990), and the differences in properties between the two suggest differences in functional role. Simple cells have strong spatial phase dependence while complex cell response is dominated by elevation of mean firing rate. If building global descriptions of contours depends on interactions between neurons with nearby RFs (for example between endstopped cells with RFs tangent to a common circle), then the position or phase-specificity of simple RFs would permit more spatio-temporally selective interactions than complex cells.

5.2 Endstopping Circuitry

Hubel and Wiesel (1965) discussed two models for generating hypercomplex cells. In one the endzones are synthesized from a single long RF neuron, and in the other the endzones derive from multiple, displaced, similarly-oriented cells. To a first approximation both models exhibit similar behaviour (orientation-dependent end inhibition) with finer scale differences depending on the precise structure of the components and the degree of nonlinearity of their combination. We have focused on the first of the two because it is the simplest model capable of explaining the principal data.

We mention two circuits for generating endstopped neurons. In one, long RF neurons in Layer 6 inhibit shorter RF neurons in Layer 4 (see McGuire et al., 1984; Bolz & Gilbert, 1986).

A second circuit is suggested by the experiments of Weyand et al (1986). In this scheme, superficial layer complex cells form the discharge region of Layer 5 endstopped special complex cells via the apical dendrite or projection to Layer 5. The endzones could also result from this latter projection or be synthesized on the laterally spreading basal dendrites in Layer 5.

Hammond and Pomfrett (1989) find approximately half of their sample of special complex cells (35 out of 68) to be nonendstopped. On the other hand, in their investigation of Layer 5 neurons Weyand et al., (1986) find the special complex property to be highly correlated with the presence of endstopping. Their striking finding was that the supragranular layers are a necessary input for the special but not the standard complex cells of Layer 5. They consider both the possibility that Layers 2-3 are necessary to drive the special complex cells, and that these layers serve to transform a standard complex cell to a special complex cell. In view of the correlation with endstopping these authors suggest that both the discharge region and endzones are supplied via the supragranular layers. As suggested above, a simpler possibility is that the discharge region is derived from supragranular inputs, but that the endzones are synthesized on the basal dendrites of the Layer 5 pyramids. This proposal has the virtue that the RF center is derived from a process (apical dendrite) or processes with less lateral extent than those which give rise to the endzones. It was observed earlier that the endstopped complex model naturally displays some of the characteristics of special complex cells. However, also note that convergence of multiple Layer 2-3 complex cells onto an *electrically short* piece of apical dendrite has the required characteristics for the subadditive synaptic mechanism discussed earlier which could underlie the special complex property independent of the presence of endstopping.

Finally, the simulations show that a complex cell whose response is the sum of input simple cell responses does not exhibit broader orientation tuning than its simple components or less length summation. However both of these will occur if the inputs are combined subadditively, and this seems to us the simplest way of understanding the observed differences in length and

orientation responses between standard and special complex cells. Hammond and Pomfrett (1989) have suggested a functional distinction to complement the mechanistic difference suggested by the complex model. They interpret the greater directionality and broader orientation tuning of special complex cells in terms of specialisation for motion rather than contour processing. Although it is reasonable to suppose that cells specialised for motion or contour analysis are biased to different operating points, unequipped with temporal dynamics, the complex model used here cannot address that issue. Also, Versavel et al. (1990) report that in addition to neurons that prefer high curvatures (strongly endstopped cells) and those preferring low curvatures (length summing nonendstopped cells) a third class of cells respond to all curvatures. Although this class was diverse in orientation tuning and degree of endstopping, many of its members appear to be special complex. This is supported by the laminar distribution found. We note that when the endstopped complex model is biased to exhibit broader orientation tuning and less length summation (within the discharge region) it naturally exhibits broader curvature response.

5.3 Endstopping and Psychophysical Curvature Discrimination

Watt and Andrews (1982) compared the curvature discrimination of human observers to an ideal statistical observer and found that the relative efficiency of the human increases with orientation range up to about 40 degrees and then diminishes. Wilson (1985) also examined curvature discrimination and compared the results to the performance of his line element model — a point processor model in which the vector of responses of a number of spatial frequency and orientation selective mechanisms is compared for different patterns. Wilson and Richards (1989) conclude that the model corresponds to the human data at high curvatures but not at low. For low curvatures they advocate pooling responses of elements along the tangent direction, a special case of cocircular tangent interactions (Parent & Zucker, 1989). A method of computing the connection compatibilities (and hence interaction strength) between discrete tangent and

curvature estimates in nearby local coordinate frames has been developed by Iverson (1988).

Wilson (1985) tested the behaviour of the line element model with curves of increasing arc length and found results essentially similar to Watt and Andrews (1982). However one curious and unexplained feature of the latter's results did not appear in the model's behaviour. In the curvature discrimination data a peak relative efficiency occurred at an orientation range of 37.5 degrees ($\kappa = 0.065$ radians/arcminute) close to the predicted value of 40 degrees, and a second, unexpected peak occurred at an orientation range of 68 degrees ($\kappa = 0.12$ rad/arcmin). Because of the number of data points these were considered to be approximately 40 and 80 degrees. The secondary peak occurred at a radius of curvature of 8 arcmin. This value is compatible with the smallest oriented foveal RFs found in primate V1 (Schiller et al., 1976; Parker & Hawken, 1988) when one takes into account that the shortest RFs are often endstopped and hence their discharge region length would be expected to be systematically underestimated (Dobbins et al., 1989). Our hypothesis is that the unexpected second peak is due to the smallest endstopped units — those that subserve the highest curvatures to which the human visual system is sensitive. This hypothesis leads to the following prediction. If one repeats the experiment in which for an arc of fixed length, curvature (and hence orientation range) are varied, at different arc lengths, the secondary peak will occur at the same curvature value, and not at the same orientation range value. If one starts with too long an arc however the smallest units will never come into play and the discrete nature of the underlying process would not be expected to emerge.

Watt and Andrews (1982) also observed that with prolonged fixation arcs of high curvature take on the appearance of a polygonal curve. This is attributable, we believe, to adaptation of the smallest or highest curvature units, leaving the remaining larger, lower curvature units to represent the curve. An indirect test involves measuring the time for the effect to set in, with and without pre-adaptation to very short lines. We predict curvature-selective adaptation with uncurved adapting stimuli.

5.4 Conclusions

In generalising the endstopped simple model to include complex components we found that the special complex property could depend on both the presence of endstopping and occlusive or subadditive interaction among inputs (abstracted in the model as saturating activation). The complex endstopped model exhibits curvature-selective responses but is not uniquely responsive to appropriately oriented and curved contours. However because of the full wave rectification of subunits in the complex model the endzone inhibition is phase independent. On the one hand this makes the inhibition more all-encompassing than if the endzones were synthesized from a simple cell, while on the other it precludes the possibility of sign of curvature selective responses arising from a spatial phase difference between the discharge region and endzones.

6. References

- Bolz, J. & Gilbert, C.D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature* **320**: 362-365.
- Desimone, R. & Schein, S.J. (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J. Neurophysiol.* **57**(3): 835-868.
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1987) Endstopping in the visual cortex as a substrate for calculating curvature. *Nature* **329**, 438-441
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1989) Endstopping and curvature. *Vision Res* **29**(10): 1371-1387.
- Dobbins, A., Zucker, S.W., & Cynader, M.S. (1990) Endstopped Visual Cortical Neurons I. Analysis of a Quasi-Linear Model. (companion paper)
- Dreher, B. (1972) Hypercomplex cells in the cat's striate cortex. *Investig. Ophthalmol.* **11**: 355-356.
- Gilbert, C.D. (1977) Laminar differences in receptive field properties of cells in cat primary visual cortex. *J. Physiol. (Lond)* **268**: 391-421.
- Hammond P., & Pomfrett, C.J.D. (1989) Direction and orientational tuning of feline striate cortical neurones: correlation with neuronal class. *Vision Res.* **29**: 653-662.
- Heggelund, P. (1981) Receptive field organization of complex cells in cat striate cortex. *Exp. Brain Res.* **42**: 99-107.
- Heggelund, P. (1986) Quantitative studies of the discharge fields of single cells in cat striate cortex. *J. Physiol. (Lond)* **373**: 277-292.
- Hoffmann, K.P. & Stone, J. (1971) Conduction velocity of afferents to cat visual cortex: a

- correlation with cortical receptive field properties. *Brain Res.* 32: 460-466.
- Hubel, D.H., & Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond)* 160: 106-154
- Hubel, D.H., & Wiesel, T.N. (1965) Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.* 28: 229-89.
- Iverson, L.A. (1988) The description of image curves: discrete forms of continuity in space. M. Eng. Thesis, McGill University.
- Jones, J.P. & Palmer, L.A. (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.* 58: 1187-1211.
- Kato, H., Bishop, P.O., & Orban, G.A. (1978) Hypercomplex and simple/complex cell classifications in cat striate cortex. *J. Neurophysiol.* 41: 1071-95.
- Koenderink, J.J. & Doorn, A.J. van (1987) Representation of local geometry in the visual system. *Biol. Cybern.* 55: 367-355.
- Koenderink, J.J. & Richards, W. (1988) Two-dimensional curvature operators. *J. Opt. Soc. Am. A* 5: 1136-1141.
- Leclerc, Y.G. (1989) Constructing simple stable descriptions for image partitioning. *Int. J. Computer Vision* 3(1): 73-102
- McGuire, B.A., Hornung, J-P., Gilbert, C.D., & Wiesel, T.N. (1984) Patterns of synaptic input to layer 4 of cat striate cortex. *J. Neurosci.* 4: 3021-33.
- Movshon, J.A., Thompson, I.D., & Tolhurst, D.J. (1978) Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol. (Lond)* 283: 79-99.
- Orban, G.A. (1984) *Neuronal Operations in the Visual Cortex*. Springer-Verlag, Berlin.
- Orban, G.A., Kato, H., Bishop, P.O. (1979a) End-zone region in receptive fields of hypercomplex

- and other striate neurons in the cat. *J. Neurophysiol.* **42**: 818-32
- Orban, G.A., Kato, H., Bishop, P.O. (1979b) Dimensions and properties of end-zone inhibitory areas in receptive fields of hypercomplex cells in cat striate cortex. *J. Neurophysiol.* **42**: 833-49.
- Parent, P. & Zucker, S.W. (1989) Trace inference, curvature consistency, and curve detection. *IEEE Trans. PAMI.* **11** (8): 823-839.
- Parker, A.J. & Hawken, M.J. (1988) Two-dimensional spatial structure of receptive fields in monkey striate cortex. *J. Opt. Soc. Am. A* **5**: 598-605.
- Rall, W. (1970) Cable properties of dendrites and effect of synapse location. In *Excitatory Synaptic Mechanisms* P.O. Andersen and J.K. Jansen, eds., pp. 175-187. Universitetsforlaget, Oslo.
- Rall, W., Burke, R.E., Smith, T.G., Nelson, P.G., & Frank, K. (1967) Dendritic location of synapses and possible mechanisms for the monosynaptic EPSP in motoneurons. *J. Neurophysiol.* **30**: 1169-1193.
- Schiller, P.H., Finlay, B.L., & Volman, S.F. (1976) Quantitative studies of single-cell properties in monkey striate cortex I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* **39**, 1288-1319, 1976.
- Schwark, H.D., Malpeli, J.G., Weyand, T.G. & Lee, C. (1986) Cat area 17. II. Response properties of infragranular layer neurons in the absence of supragranular layer activity. *J. Neurophysiol.* **56**(4): 1074-1087.
- Singer, W., Tretter, F., & Cynader, M. (1975) Organization of cat striate cortex: a correlation of receptive field properties with afferent and efferent connections. *J. Neurophysiol.* **38**: 1080-1098.

- Spitzer, H. & Hochstein, S. (1985) A complex-cell receptive field model. *J. Neurophysiol.* **53**: 1281-1301.
- Tanaka, K., Ohzawa, I., Ramoa, A.S., & Freeman, R.D. (1987) Receptive field properties of cells in area 19 of the cat. *Exp. Brain Res.* **65**: 549-558.
- Versavel, M., Orban, G.A. & Lagne, L. (1990) Responses of Visual Cortical Neurons to Curved Stimuli and Chevrons. *Vision Res.* **30**: 235-248.
- Watt, R.J. & Andrews, D.P. (1982) Contour curvature analysis: Hyperacuities in the discrimination of detailed shape. *Vision Res.* **22**: 449-460.
- Weyand, T.G., Malpeli, J.G., Lee, C., & Schwarkj, H.D. (1986) Cat area 17. III. Response properties and orientation anisotropies of corticotectal cells. *J. Neurophysiol.* **56**(4): 1088-1101.
- Wilson, H.R. (1985) Discrimination of contour curvature: data and theory. *J. Opt. Soc. Am. A* **2**: 1191-1199.
- Wilson, H.R. & Richards, W. (1989) Mechanisms of contour curvature discrimination. *J. Opt. Soc. Am. A* **6**: 106-115.
- Zucker, S.W., Dobbins, A., & Iverson, L. (1989) Two stages of curve detection suggest two styles of neural computation. *Neural Computation* **1**: 68-81.

7. Appendix 1.

The positive part function $\Phi(R; \rho, \Gamma)$ is closely related to elementary functions. When $\Gamma = 1$, it reduces to a hyperbolic tangent, which in turn can be written as a difference of standard and reversed logistic functions:

$$\Phi(R; \rho, \Gamma = 1) = \tanh(2R/\rho) = \frac{1}{1 + e^{-R/\rho}} - \frac{1}{1 + e^{R/\rho}}. \quad (A1)$$

Re-introducing Γ and fixing $\Gamma = \frac{1}{e^{R_0/\rho}}$, the positive part function can be written as,

$$\Phi(R; \rho, \Gamma) = \frac{1}{1 + e^{-(R-R_0)/\rho}} - \frac{1}{e^{R/\rho} + e^{R_0/\rho}}. \quad (A2)$$

where the first term is a sliding logistic function, and the second term determines the amount of compression of negative values. This form emphasizes that gain (ρ) and spontaneous firing (Γ) are not decoupled (see also Figure 2).

8. Figure Legends

Figure 1. Models of the simple and complex cell components. Spatial integration is followed by a saturating, rectifying nonlinearity.

Figure 2. The generalized rectification function $\phi(R; \rho, \Gamma)$. R is the input and as Γ (written S for spontaneous in figure) varies from 0 to 1 the output goes from half wave rectified to unrectified. ρ controls the contrast sensitivity or gain.

Figure 3. The response of a complex cell instance and one of its simple components as a function of line position, orientation and width. The simple cell parameters are described by an even symmetric Gabor RF with length 61 pixels, envelope aspect ratio (AR) 2.0, and modulation frequency $u_0 = 0.05$ (5 pixels = 1/4 wavelength). The complex instance has 3 simple components with spacing 3 pixels and weightings given by a Gaussian with $\sigma = 4$. The line weighting function (bar response as a function of position) is shown for long bars and widths of 3, 5, and 9 pixels (solid, dotted and dashed curves, respectively). The fourth complex cell line weighting function (very long dashes) illustrates half wave rectification of subunit responses (line width 9). Notice the apparent reduction in RF width that results. The orientation response is for bars centered on the RF and rotated in steps of 4 degrees for the same bar widths.

Figure 4. Length, orientation and curvature tuning curves for a complex model instance biased for subadditivity. The dotted ($\rho = 65$) and dashed ($\rho = 43$) curves illustrate two degrees of saturation while the solid curve repeats the nonsaturating characteristic of the previous figure. For convenience of comparison all outputs are normalized.

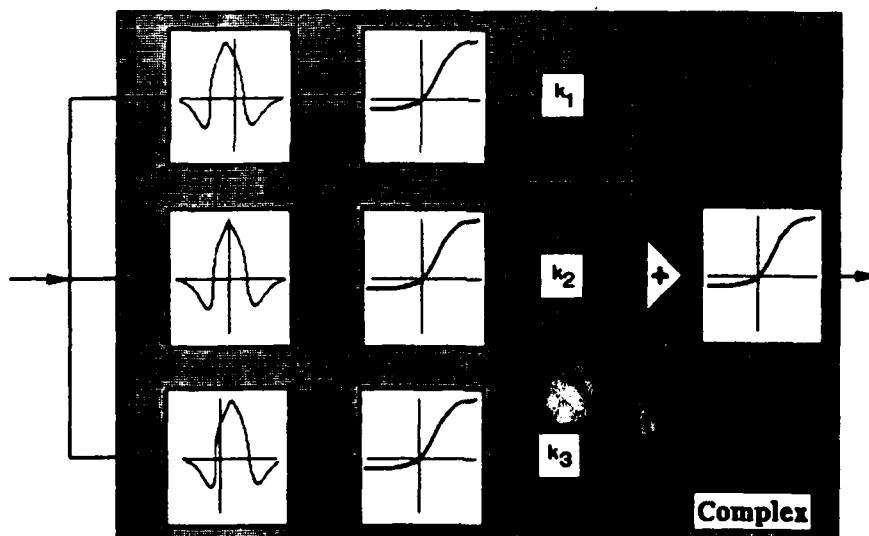
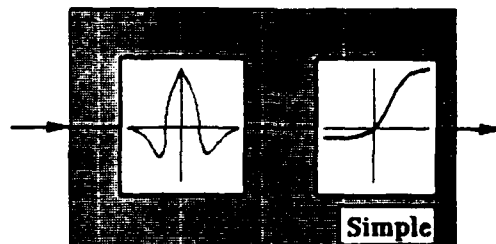
Figure 5. The response of an endstopped instance in which the small cell is simple (size:35, AR:2.0, $u_0=0.05$, $c^S=2.8$) and the large cell is complex (size:61, 3 components with spacing=4, and Gaussian weighting $\sigma = 10.0$. Each component has AR=3.0 and $u_0=0.05$.) a. The response to (binary) curves varying in radius from 0 to 200 in steps of 10 as a function of curve width. The curves are of width 1,3,5,7, and 9 pixels (solid, dotted, dashed, long dashed, and dot-dashed,

respectively). b. The response to curves of orientations from 0 to 40 degrees from optimal in steps of 10 (curve width 5). The solid curve corresponds to the dashed curve in part a.

Figure 6. The response of an endstopped instance in which both small and large component cells are complex. Both the small and large complex cells are composed of 3 simple components of size 35, spacing 4. The small and large cells have Gaussian weighting $\sigma = 4.0$ and 10.0, gains $c^S = 1.9$ and $c^L = 1.0$, and aspect ratios of 1.0 and 3.0, respectively. a. The response to (binary) curves varying in radius from 0 to 200 in steps of 10 as a function of curve width (solid, dotted, and dashed are of width 1, 5, and 9 pixels respectively). b. The response to curves of orientations from 0 to 40 degrees from optimal in steps of 10 (curve width 5). The solid curve corresponds to the dashed curve in part a. c. The response to a curved line (width 5 pixels, solid curve) and curved edge (dotted line) of radius 30 pixels as a function of position across the RF. At the leftmost edge of the graph the curve is at the boundary of the RF and curves away from the RF, at the center it is centered on the RF and at the rightmost edge is on the boundary of the RF curving toward the RF. d. The curvature response to a curved edge when the edge is centered on the RF. The curve orientation is varied, and the response curves for orientation deviations 0 and 10 degrees are indistinguishable, while there is no response to 20 or 30 degree deviation.

Figure 7. a. The length response for the endstopped complex cell of the previous figure (line width 5, step size 2). b. The orientation response of the same cell for lines corresponding to the peak of the length-tuning curve (22 pixels, solid curve), the length of the excitatory center (35 pixels, dotted curve), and longer than the discharge region plus endzones (80 pixels, dashed curve). c. The orientation response when the positive part function is employed with moderate spontaneous firing ($\Gamma = 0.5, \rho = 40$) for the excitatory component (solid, dotted, and dashed curves show the response of the endstopped cell and its small and large components, respectively). d. The curvature tuning curve for the cell (with spontaneous firing as in part c.) as

curve orientation is perturbed (0,10,20,30 degrees for the solid, dotted, dashed and long dashed curves, respectively).



b.

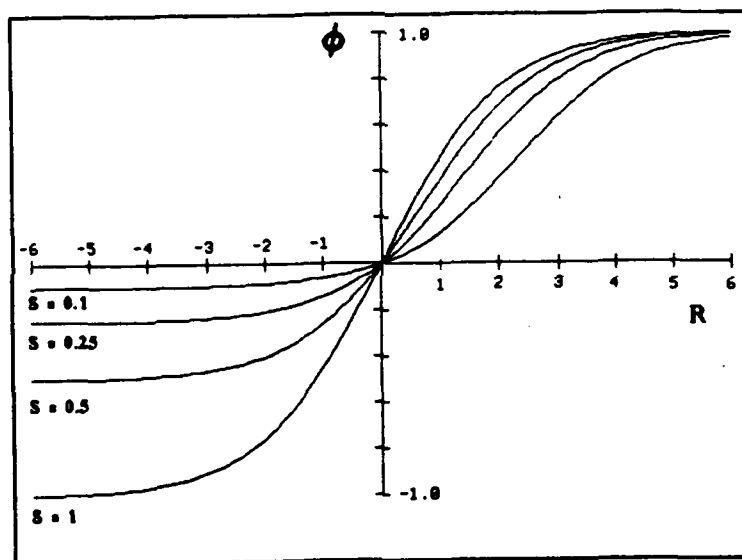


Figure 2

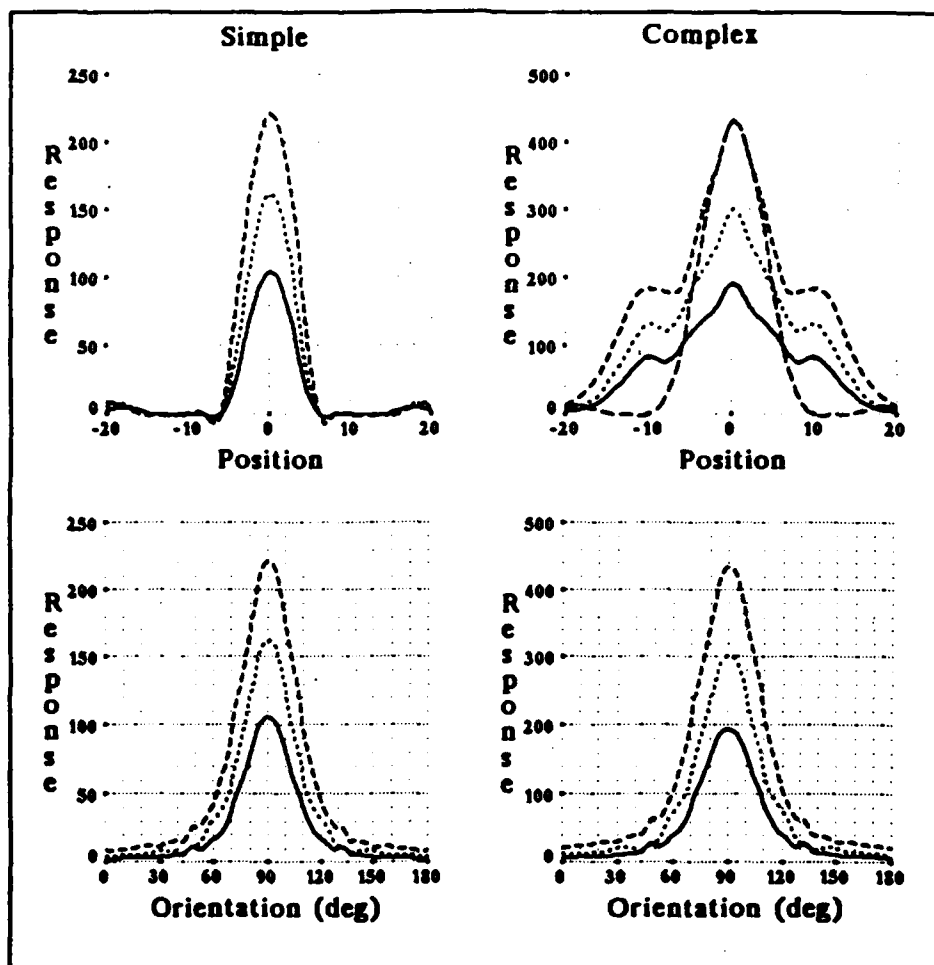
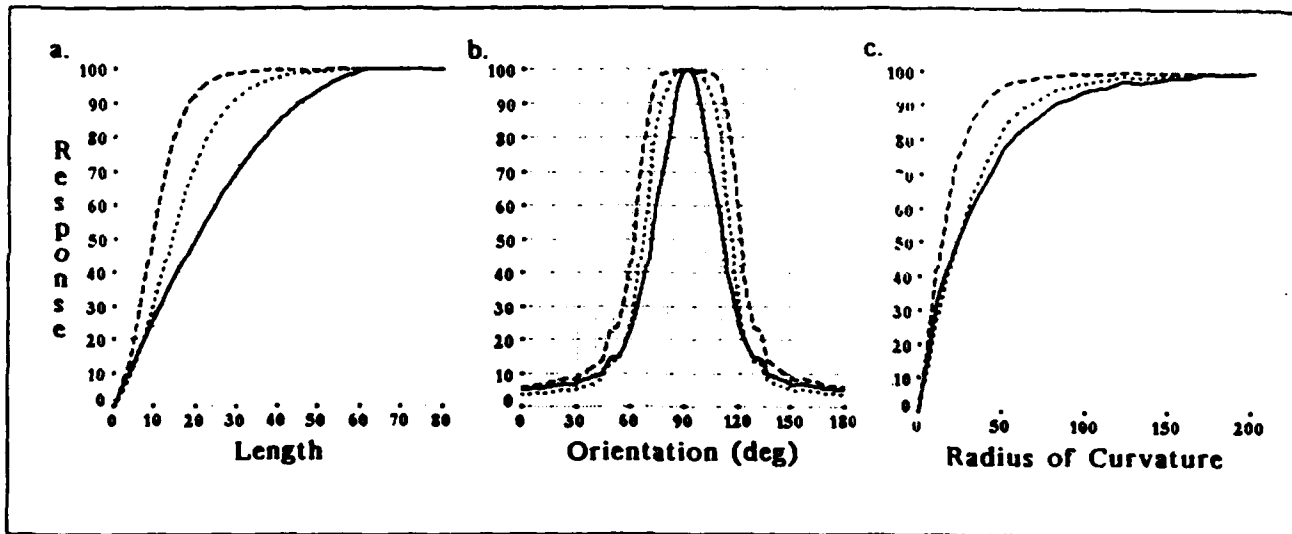
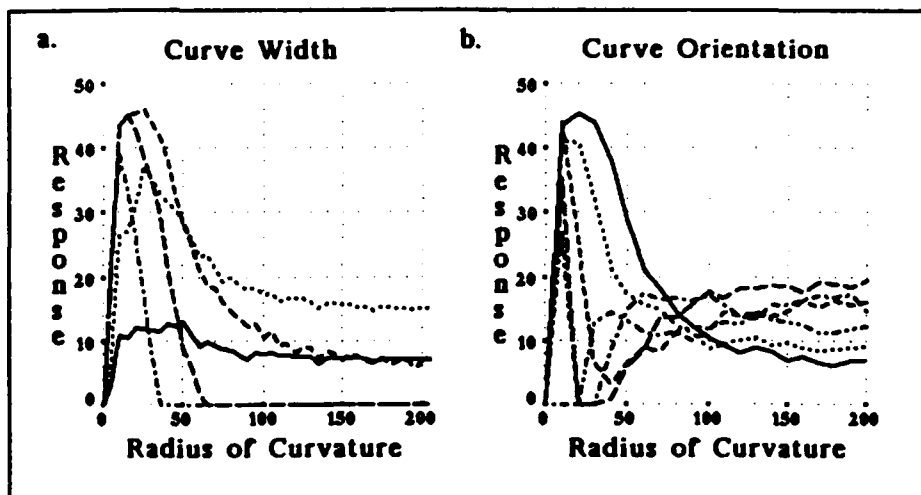


Figure 3





7005

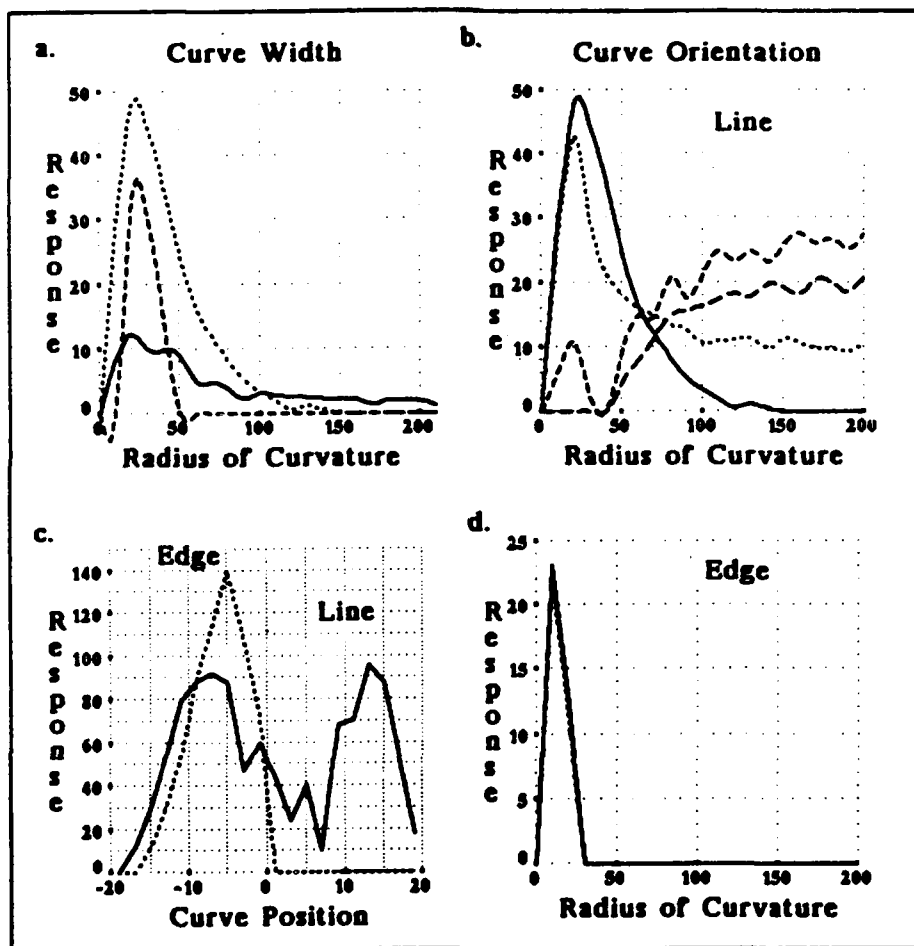
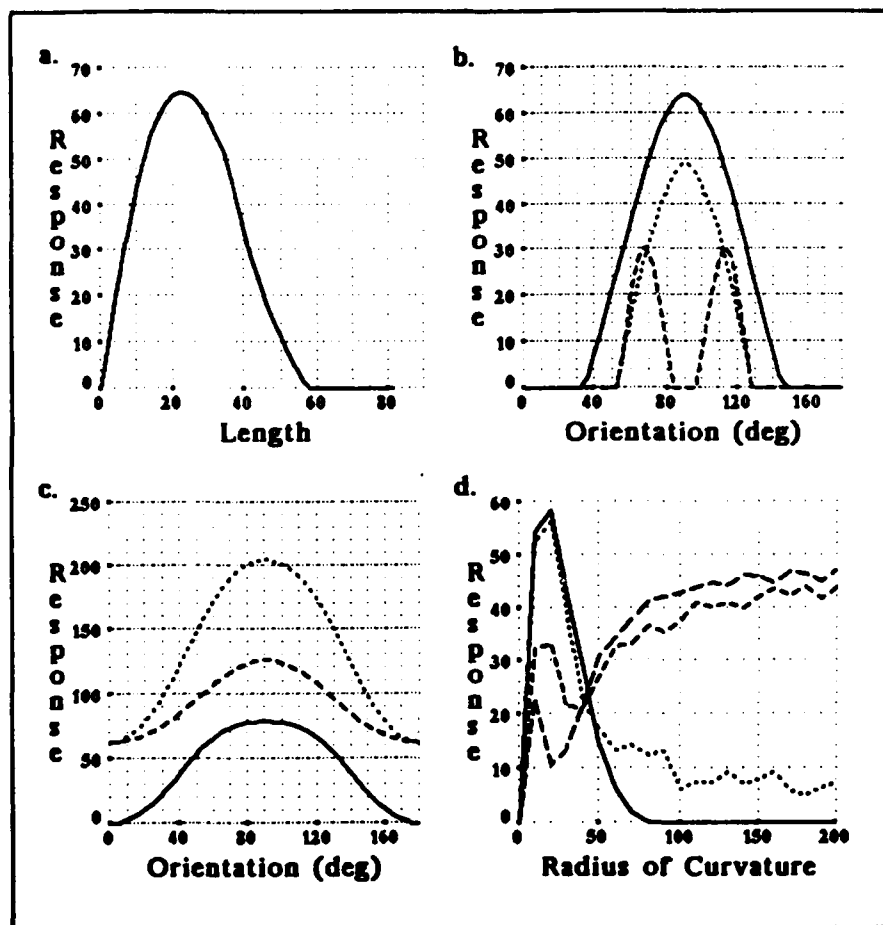


Figure 6



Appendix V

A Mean Field Model of Optic Flow Estimation, 1991 ARVO abstract, and manuscript.

DO NOT FOLD THIS FORM

1991 ARVO Abstract Form

Numbers on this form refer to
Instructions on pages 2-3

PRESENTATION PREFERENCE

Please check one (1).

- ☐ (POS) Poster only or Withdraw
☐ (NP) Paper or Poster
(no preference)
☐ (PA1) Paper #1, Poster #2
☐ (POS1) Poster #1, Paper #2

SCIENTIFIC SECTION PREFERENCE

See Section Descriptions on page 16
Please check the one (1) Section
best suited to your proposed
presentation

- ☐ (A) Anatomy/Pathology
☐ (B) Biochemistry/Molecular Biology
☐ (C) Clinical/Epidemiologic Research
☐ (D) Cornea
☒ (E) Electrophysiology
☐ (F) Eye Movements/Strabismus/
Amblyopia
☐ (G) Glaucoma
☐ (I) Immunology/Microbiology
☐ (L) Lens
☐ (P) Physiology/Pharmacology
☐ (R) Retina
☐ (T) Retinal Cell Biology
☐ (V) Visual Psychophysics/
Physiological Optics

TOPIC CODE

Write the number of one
(1) Topic Code that most
accurately reflects the
content of your proposed
presentation.
See pages 4-5

161

TRAVEL FELLOWSHIP GRANT

Put a check in the box if
you are applying for a
Travel Fellowship Grant.

☐

ABSTRACT PREPARATION

Type title
of the abstract, authors'
names, the institutions'
names, and the actual
abstract within the
borders of the blue box to
the right. See specific
instructions on page 3.

Abstract Forms must be
received at the ARVO
Central Office by 5:00 p.m.,
Friday, December 7, 1990.

**COPYRIGHT
TRANSFER SECTION:**
Complete section on
reverse side.

FIRST (PRESENTING) AUTHOR

- ☐ 1991 ARVO Member
(Membership Number—Look on Mailing Label)
☐ Non-Member (must have Sponsor—see ☒ below)
☒ Applying 1991 ARVO Member (Membership Application Form enclosed)

Last: Dobbins First: Allan Middle: C.
Department: Electrical Engineering Institution: McGill University
3480 University, Montreal Quebec CANADA H3A 2A7
Street Address:
City: Montreal State/Country: Quebec Zip + 4: H3A 2A7 Province:
Office Phone: (514) 398-7153 Home Phone: (514) 849-9360

SPONSOR (Must be a 1991 ARVO member if First Author is not; cannot be First Author on another abstract; and must be a co-author on this abstract.)

Zucker, Steven W.
Last: Zucker First: Steven Middle: W.
Office Phone: (514) 398-7134 Home Phone: (514) 933-8520
Membership Number: 35878

A MEAN FIELD MODEL OF OPTIC FLOW ESTIMATION

Allan Dobbins† Steven W. Zuckert‡ Max S. Cynader§
†McGill University, Montreal §University of British Columbia, Vancouver.

We address two questions. How can the visual system obtain reliable estimates of the optic flow field? How can local motion and disparity estimates be related to the shape of visible surfaces? From the observed properties of visual cortical cells we build model neurons and local circuits that are evaluated in dynamic simulations. Neurons in the middle temporal visual area (MT) are direction and speed selective and many also have direction and speed-dependent antagonism between the receptive field (RF) center and surround. Given an ensemble of such cells with RFs spanning a range of sizes, it is possible in principle to obtain maximal mean field approximations of the optic flow field — the largest neighbourhoods around each point in which the field is well-approximated as uniform translation. The advantage of this approach is that estimation of the rate of spatial variation of the field yields improved signal-to-noise ratio without blurring across distinct spatial regions. Although center-surround speed bandpass neurons should respond well to spatially restricted translation, or to appropriately curving, accelerating, or discontinuous flows; they provide no information about which of these situations is, in fact, the case. Estimation of the large scale first order variation of the flow field appears to occur subsequently (Saito *et al*, 1986), however it is the second order variation that is related to surface shape. We shall show how neurons may be organized to provide coarse information about the shape of surfaces from the second order variation of the optic flow field and from binocular disparity.

Mean Field Theory and MT Neurons

Allan Dobbins[†] Steven W. Zucker^{†*}

[†] Computer Vision and Robotics Laboratory
Department of Electrical Engineering
McGill University
Montréal, Québec, Canada

* Senior Fellow, Canadian Institute for Advanced Research.

Abstract

In this paper we consider the problem of obtaining reliable estimates of the optic flow field and its variation. We shall argue that circuitry in the middle temporal visual area (MT) is arranged to provide mean field approximations of the optical flow field — the maximal neighbourhood over which the local field can be well-approximated as uniform, parallel translation. In addition, the way the mean field approximation varies with neighbourhood size and position is informative about the variation of the flow field. Further, we show how receptive fields can be synthesized to represent the first order differential information about the flow field. The same ideas are applicable to the binocular disparity vector field, and lead us to predict the existence of extrastriate neurons that are specialised for the local shear and stretch of the disparity and motion vector fields.

1. Introduction

Our goal is to understand how the brain estimates surface shape from the visual information available to it. This problem had played a central role in theories of vision [Gibson, 1950; Marr, 1982]. Here we shall consider those cases in which there are vector fields associated with the surfaces: the motion flow field and the binocular disparity field.

Koenderink has repeatedly emphasized the value of the differential structure of vector fields. In the optic flow field these are invariant with a constant translation whether due to object movement or an eye movement. Thus the appropriate setting for the first order differential information is affine space. In both the binocular disparity field and the motion flow field, all the information about surface orientation is contained in the deformation component [Koenderink & van Doorn, 1975, 1976; Koenderink, 1986].

In this paper we shall show how the brain could estimate components of the gradient of the optic flow field. We shall argue that instead of simply differentiating the flow field and combining the partial derivatives, the circuitry of the middle temporal cortex (MT) determines maximal mean field estimates — the largest convex regions where the flow is well-approximated as parallel. If such patches can be obtained, certain simplifications follow in the gradient computations.

In the next section we treat the generic features of visual cortical architecture that we think underlie the computations that are described subsequently.

2. Visual Cortex: Estimators and Architecture

In this section we briefly review some features of the organization of the primary visual cortex (V1) and the middle temporal area (MT) that are germane to the developments that follow.

Primary visual cortex is organized into columns of cells which share their preferred orientation [Hubel & Wiesel, 1962]. Thus each retinal point is multiply represented in a set of columns spanning orientation space. In MT, an area which a variety of evidence indicates is specialised for motion analysis, the cortex is also organized into columns [Albright et al. 1984]. In this case the units in a column share a preference for direction of motion. In addition, MT neurons are speed-selective with individual cells exhibiting a broad range of preferred speeds and bandwidths (Maunsell & Van Essen, 1983).

Some of the cells in V1 exhibit an orientation-specific, antagonistic relationship between the receptive field center and zones beyond the ends of the excitatory receptive field. Originally termed hypercomplex when they were first described in extrastriate cortex [Hubel and Wiesel, 1965], these neurons exhibit a differential response with stimulus length, and are commonly referred to as endstopped. We have previously described a computational model of endstopped cells which is length selective, but which is curvature-selective in addition. Hubel and Wiesel recognized that endstopped cells should respond well to highly curved stimuli that avoid the endzones, and recent neurophysiological experiments demonstrate that endstopped cells are curvature-selective [Dobbins et al. 1987]. In area MT many cells exhibit a direction and speed specific antagonism between center and surround [Allman et al. 1985]. They also exhibit differential response with flow field variation. In the following section we develop a view of the estimation carried out by MT circuitry.

3. The Mean Field Model

Although field theories can be formulated in differential (point) or integral form, it makes little sense for computa-

tional vision theories to be formulated in differential form. There are several reasons. One is that because of the uncertainty relation between a signal and its derivative, one cannot obtain information about the infinitesimal, differential information in a flow field. A second constraint is imposed by the representational precision of the local estimators: shrinking the temporal neighbourhood shrinks the data precision correspondingly. Thirdly, the surface features in motion can change in size, orientation, and shape. For example, if we modeled the time-variation of an image feature as a wave group, its motion is characterised by both phase (ω/k) and group ($d\omega/dk$) velocities.

Since the visual system cannot have access to the point information, it must employ spatio-temporal integral measures. An essential part of its task is determining the neighbourhood size over which a particular order estimate holds. Here we propose an approach to how the part of the cortex specialised for motion analysis determines the best linear approximation to the flow over various neighborhood sizes.

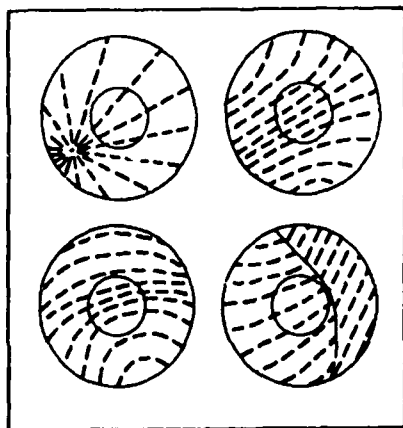


Figure 1 Four flow patterns, one of which is discontinuous, all of which are well-approximated by uniform translation over the small neighbourhood disk but not over the large one.

3.1 Definition of Terms

By the optic flow field we shall mean the mapping, $v: R^2 \rightarrow R^2$, from retinal coordinates to TM the tangent bundle of the flow field. Hence the image of x at time t_0 is the velocity vector $v(x)|_{t=t_0}$. The direction field is a mapping from retinal coordinates to the direction of the flow velocity ($x \mapsto v/\|v\|$). By a mean field approximation of the flow field we shall mean the vector or differential 1-form that is the estimated average translation velocity over a specified disk ($D(r, x)$) around a point. Naturally, this comprises both the average speed and direction.

3.2 Initial Measurements

Our goal is to estimate the affine, linear structure of the flow field. Toward this end we assume the availability of two types of local, speed selective, direction selective units.

broadband units These are cells which have a speed bandwidth of two to three decades (r_{θ_i}).

narrowband units These are cells which have a speed bandwidth of two to four octaves (r_{θ_i, v_j}).

The narrowband units span the range of speeds to which the system is sensitive, and both types of units cover the set of retinal directions and the retina itself. We shall not be concerned here with detailed description of these cells, but mention that they are meant to correspond to Area V1 and V2 units.

3.3 Pooling Responses

At the next stage the responses of many of these units are pooled, and it is to the function of this process that we now turn. Let us sum over a convex retinotopically-mapped region, the responses of a set of narrowband units whose receptive fields cover the region. All the units are tuned to the same direction (θ_i) and speed of movement (v_j). Although the image and inputs are discrete, for convenience we shall refer to integration. The spatial integral of the unit responses over a disk $D(r, x)$ is:

$$R_{\theta_i, v_j}^{D(r)}(v, x) = \int_D r_{\theta_i, v_j}(v, x) dA. \quad (1)$$

This measures the degree of correspondence of the flow field to a uniform, parallel flow in the direction to which the units are tuned. Normalizing by area, we obtain a space-averaged measure of the projection over the retinal area in question:

$$\bar{R}_{\theta_i, v_j}^{D(r)}(v, x) = \frac{1}{A} R_{\theta_i, v_j}^{D(r)}(v, x). \quad (2)$$

Alternatively, summing over the broadband units (r_{θ_i}), one obtains a measure that varies with direction but not with speed. Therefore a pattern translating in the right direction but with inhomogeneous speed will elicit strong response.

We noted above that the distribution of response with direction contains important information about the field structure, but it also should be clear that the set of directional spatial averages over a particular neighbourhood does not uniquely determine the flow field (see Figure 1). Therefore we investigate the significance of the variation in the mean field estimates with integration area.

3.4 Analysis of Flow Field Variation with Area

Now we wish to examine how the integrated response varies with area. For convenience we shall simplify the notation, ignoring the dependence of R on position and the flow field, and simply refer to its area-dependence $R(A)$. Either R or \bar{R} can be employed for the analysis. It is assumed that the local unit measures are nonnegative, implying that $\delta R \geq 0$. **Case 1.** $R(A)$ constant. a. If $R(A) = 0$, then there is no component of the flow field in the direction θ . b. If $R(A) > 0$, then $\bar{R} \rightarrow 0$ as $1/A$. For the minimal disk of area A_0 there

is a response but there is no contribution from the surrounding area. Thus there is either no field outside A_0 or it is orthogonal to θ .

Case 2. $R(A)$ increasing. There is a contribution to the integral outside A_0 . There are three principal cases according to the sign of the second derivative of R with respect to A .

a. $d^2R/dA^2 > 0$. R inflects up implying that the field corresponds better to θ over a larger area. This occurs when the average field moves toward the estimators either in direction or speed. Generally, θ would not be the mean field estimate at smaller areas. Thus this case generally represents a change in the mean field approximation with area.

b. $d^2R/dA^2 = 0$. R increases proportional to area. This indicates that the field is not changing systematically with respect to θ . It does not necessarily imply that θ is the best mean field estimator however. For example, if the field is a uniform field translating at an acute angle with respect to direction θ , R increases linearly with A , the slope depending on the relative angle. In this case there should be another direction that is the best mean field estimate at all areas for which this condition holds.

c. $d^2R/dA^2 < 0$. R inflects downward, implying that the estimators conform better to the field over smaller areas. If θ is the mean field direction for small A but not large A then the field direction is changing (either smoothly or abruptly). If θ is the mean field direction for both small and large A , then the field is either less parallel (the distribution of R over the set of directions is broadening with increasing A), or speed is changing.

Note that we did not specify whether integration was taking place over narrowband units, broadband units or both. Suppose that separate integrations are done for the narrowband units tuned to separate speeds, as well as for the broadband units. Since the latter are concerned only with direction and not with speed (over a broad range) their change in response can only be due to a change in direction. In contrast the narrowband unit integration changes with both speed and direction. However the difference between the change in response due to direction and speed, and that due to direction, is the change in response due to speed change. Assuming quasi-orthogonality (e.g. a stimulus halfway between centers of two adjacent direction or speed classes elicits a half response from each), summing over speed and direction classes, and subtracting broad from narrowband totals, one could in principle partition the variation into speed and directional variation. We note in passing that the *convexity* detector model of Nakayama and Loomis [1974] involved summing (center-surround velocity filters) over direction to detect speed discontinuities in the optic flow field. The interpretation of the responses of such detectors is problematic in a world with nonplanar surfaces however.

3.5 MT Receptive Field Organization

Now we shall attempt to connect the process described above to what is known of cortical organization. We are interested in the value of $R(A)$ and its first two derivatives.

However it is easily shown that this information is contained in the first variation of \bar{R} . To see this, consider the difference in the value of \bar{R} over disks of area a and b ($b > a$). Some rearrangement yields:

$$\bar{R}(b) - \bar{R}(a) = \frac{R(b)}{b} - \frac{R(a)}{a} = \frac{b-a}{ab} \left[\frac{R(b) - R(a)}{b-a} - \frac{R(a)}{a} \right] \quad (3)$$

The terms in braces are the average projections of the flow field onto the annulus of outer radius b and the disk of radius a . Therefore the whole expression is the difference of the field's projection onto the annulus and the disk it surrounds. Hence instead of evaluating the infinitesimal second derivative of R , the finite first variation of the space-average measure is used.

Consider a finite difference scheme in which $\bar{R}(A)$ is estimated at intervals e.g., $A_0, 2A_0, 4A_0, \dots$. Then the difference of the successive area-normalized integrals is computed: $\bar{R}(A_0) - \bar{R}(2A_0)$ and so forth. In physical terms, we suppose that there is a set of neurons the responses of which correspond to $\bar{R}(A_0)$, $\bar{R}(2A_0)$ and so forth, and further, that are neurons which represent the differences of successive terms in this sequence. These neurons encode the variation of the flow field over successive area increments. In addition to capturing the rate of change of the flow field, this circuitry can also be exploited to determine the largest neighbourhood over which the field is well-approximated by parallel translation. In particular if θ is the direction of maximum response over a disk A_i and $\bar{R}_\theta(A_i) - \bar{R}_\theta(A_{i+1}) \leq 0$, then A_i can be replaced in subsequent computations by A_{i+1} .

It is natural to suppose that the difference computation occurs *within* a cortical column. For most smoothly varying patterns, all receptive field centers less than some size are strongly inhibited by their surrounds. Then at some point in the sequence there is a center which is both strongly activated and not strongly inhibited. The area of the center of this cell is the maximal patch area.

Note that although the magnitude of the variation of the flow field surrounding a point can be estimated in this way, this scheme does not represent the nature of the variation. There would seem to be two possibilities. The first is that the form of the variation is distributed over the set of direction and speed integrals over patches of the field. Alternatively, particular forms of variation may be explicitly represented in the response of individual neurons. In the next section we shall consider how this might be accomplished.

4. Estimating Differential Flow Structure

In this section we consider how neurons could estimate the spatial gradient of the optic flow field. The circuitry described in the last section is used to determine the largest neighbourhood in which the flow can be treated as parallel. The mean field circuitry imposes a choice of coordinate system (p, n) , representing the parallel and normal components of the flow, $v = (v_p, v_n)$, and selects a neighbourhood size

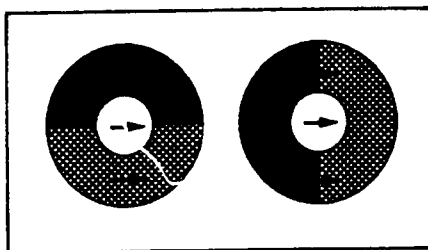


Figure 2 A schematic representation of two direction-selective receptive fields with antagonistic center and surround. The surround is split into two subzones in each case.

over which the normal component is negligible. Hence over region A, $v \approx (v_p, 0)$.

Consider the first order variation of the flow with respect to v_p . The variation is decomposed into $\frac{\partial v_p}{\partial p}$ and $\frac{\partial v_p}{\partial n}$. However, keep in mind that the estimates are always integrals over area. We now propose that these quantities can be estimated with center-surround mechanisms differing in only one respect from those described in the last section. Suppose that the surround is divided into halves. In one half the constituent units are low pass for speed and in the other half high pass. Figure 2 shows examples in which the division is parallel and perpendicular to the directional preferences of the center and surround. The additional requirement is that the rolloff rates of the low and high pass mechanisms are equal and opposite. Before proceeding, it should be mentioned that the surround units can actually be bandpass, the important thing is that they are lowpass and highpass in the window of speeds to which the center is sensitive. Figure 3a shows the arrangement. All the units (both lowpass and highpass) comprising the surround are summed as before, and the area-normalized value is subtracted from the value for the center.

Now consider the response of this mechanism to flow patterns. For a uniform, translating pattern, the cell behaves as before – the surround cancels the response of the center. This is because the equal slope condition ensures that the sum of the low and highpass units is invariant over a range of speeds. However, its response is quite different to patterns with spatial variation of speed. Suppose that, in the coordinate frame (p, n) , the velocity field is given by: $v = (v_0 + mp, 0)$ where m is the spatial gradient of the speed. The cell will respond proportional to m . To see why, note that we could equally have described the surround organization as a difference of two spatially segregated low pass (or high pass) mechanisms of identical rolloff. Thus the surround represents a first spatial difference of speed, and responds proportional to the speed gradient. This is shown in Figure 3b. The center serves to establish that the field is sufficiently parallel and has average speed within some range, and the surround modulates it proportional to the gradient. Observe that swapping the low and high pass regions reverses the sign of gradient to which the cell responds.

A cell with the surround split as above is interested in the speed change in the direction of the flow. If the surround

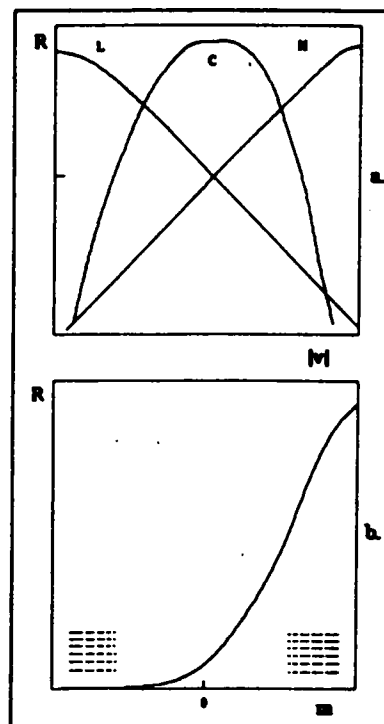


Figure 3 a. The speed dependence of the response of the center (C), and the two halves of the surround (L and H) of a hypothetical center-surround cell. b. The response of the cell with the gradient of speed, in the direction normal to the axis of division of the surround. These are not quantitative simulations, merely a qualitative sketch.

is divided parallel to the preferred direction, the cell responds proportional to the (parallel component of) speed change normal to the flow direction, that is, proportional to the shear of the field. Generally, the shear depends on $\frac{\partial v_n}{\partial p}$ as well. To the extent that the mean field circuitry has a chosen a neighbourhood size such that $v_n = 0$ the contribution of this term is negligible. Simulations are in order to determine the performance limitations of this scheme.

5. Summary and Conclusions

1. Integration of the broadband units underlies the computation of the mean direction field, and the differential of this field describes the rate of directional change with area.

2. In contrast, integration of the narrowband directional units underlies estimation of the mean velocity field, and its differential describes the rate of change of velocity (both direction and speed) with area.

3. The result of the mean field estimation process is a set of patches of varying size covering the flow field. Each patch has an associated velocity vector.

4. To the degree that the mean field circuitry succeeds in choosing a neighbourhood size such that the normal component of velocity is negligible over the neighbourhood, the vector field is locally reduced to a scalar field, and the gradient of the field is reduced from a second order tensor to a first

order tensor (a vector). The consequence is a reduction in complexity in the sense of the number of numbers associated with an estimate.

5. Although space does not permit its development here, it is possible to develop estimates of the components of the disparity gradient using opposite slope estimators (*near* and *far* cells), in an analogous manner to that described here for the components of the optic flow gradient.

6. The ideas described here have so far neither had their quantitative plausibility established in simulations, nor their truth refuted in physiological experiments. Both courses of action will be pursued.

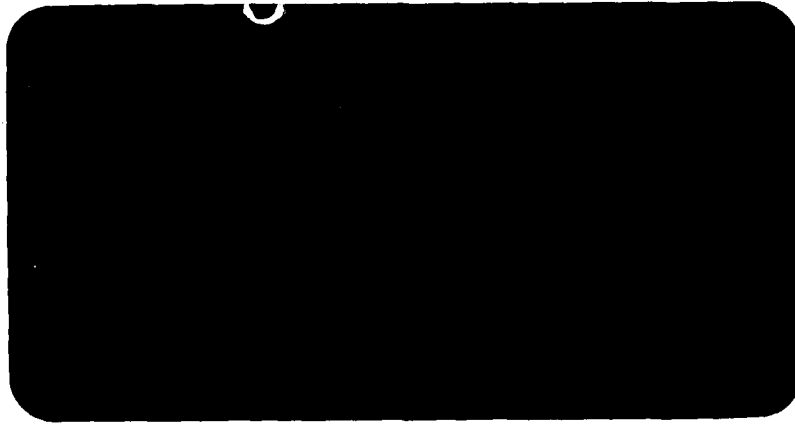
6. References

- [1] J.J. Gibson "The Structure of the Visual World". Houghton, Mifflin. Boston, 1950.
- [2] D. Marr "Vision". Freeman, San Francisco, 1982.
- [3] J.J. Koenderink & A.J. van Doorn "Invariant Properties of the Motion Parallax Field due to Movements of Rigid Bodies Relative to an Observer". *Optica Acta*, 22, 773-791, 1952.
- [4] J.J. Koenderink & A.J. van Doorn "Geometry of Binocular Vision and a Model for Stereopsis". *Biol. Cybernetics*, 21, 29-35, 1976.
- [5] J.J. Koenderink "Optic Flow". *Vision Res.*, 26, 161-180, 1986.
- [6] D.H. Hubel & T.N. Wiesel "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". *J. Physiol. London*, 160, 106-154, 1962.
- [7] T.L. Albright, R. Desimone, & C. Gross "Columnar Organization of Directionally Selective Cells in Visual Area MT of the Macaque". *J. Neurophysiol.*, 51, 16-31, 1984.
- [8] J.H.R. Maunsell & D.C. Van Essen "Functional Properties of Neurons in Middle Temporal Visual Area of the Macaque Monkey. I. Selectivity for Stimulus Direction, Speed, and Orientation". *J. Neurophysiol.*, 49, 1127-1147, 1983.
- [9] D.H. Hubel & T.N. Wiesel "Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat". *J. Neurophysiol.*, 28: 229-89, 1965.
- [10] A. Dobbins, S.W. Zucker, & M.S. Cynader, M.S. "End-stopping in the visual cortex as a substrate for calculating curvature". *Nature*, 329, 438-441, 1987.
- [11] J. Allman, F. Miezin, & E. McGuinness "Direction- and Velocity-Specific Responses from Beyond the Classical Receptive Field in the Middle Temporal Visual Area (MT)". *Perception*, 14, 105-126, 1985.
- [12] K. Nakayama & J.M. Loomis "Optical Velocity Patterns, Velocity-Selective Neurons, and Space Perception: a Hypothesis". *Perception* 3, 63-80, 1974.

Appendix VI

On the evolution of curves via a function of curvature, 1: The classical case, *J. Mathematical Analysis and Application*, to be published.

McGill Research Centre for Intelligent Machines



McGill University



**On the Evolution of Curves via a Function of
Curvature, I:
The Classical Case**

Benjamin B. Kimia * Allen Tannenbaum †
Steven W. Zucker * ‡

TR-CIM-90-3 rev: 10 Sept 1990

Computer Vision and Robotics Laboratory
McGill Research Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada

*McGill University, Dept. of Electrical Engineering

†Department of Electrical Engineering, Technion, City-Haifa 32 000 Israel
and Dept. of Elect. Eng., University of Minnesota, Minneapolis, MN
55455

‡Fellow, Canadian Institute for Advanced Research

Postal Address: 3480 University Street, Montréal, Québec, Canada H3A 2A7
Telephone: (514) 398-6319 Telex: 05 268510 FAX: (514) 283-7897
Network Address: mrcim@larry.mrcim.mcgill.edu

On the Evolution of Curves via a Function of Curvature, I: The Classical Case

Benjamin B. Kimia

Allen Tannenbaum

Steven W. Zucker

Abstract

The problem of curve evolution as a function of its local geometry arises naturally in many physical applications. A special case of this problem is the curve shortening problem which has been extensively studied. Here, we consider the general problem and prove an existence theorem for the classical solution. The main theorem rests on lemmas that bound the evolution of length, curvature and how far the curve can travel.

Keywords: Curve Evolution, Deformation, Curve Shortening, Flame Propagation, Curvature Bounds.

Acknowledgements

This research was supported by grants from NSERC, MRC, FCAR, AFOSR and NSF.

1. Introduction

In this paper, we study certain evolution equations of embedded plane curves where the speed of the deformation is a function of the curvature, and whose direction is in the normal direction.

The curve evolution problem is relevant in applied sciences. The study of immersed closed curves evolving as functions of their curvature has been carried out for crystal growth [Langer, 1980; Ben-Jacob *et al.*, 1983], flame propagation [Sethian, 1985; Sivashinsky, 1977; Osher and Sethian, 1988], and curve shortening [Gage and Hamilton, 1986; Grayson, 1987]. We would like to investigate properties of the classical solutions of these evolution equations. As pointed out by the referee, similar work has been done independently by [Angenent, 1989]. Finally the equations we use for the curvature of a plane curve evolving according to the general law given below are classical; see e.g. [Brower *et al.*, 1984]. In a sequel we plan to consider the weak solutions when shocks develop [Lax, 1973; Lax, 1971; Osher and Sethian, 1988].

The research has been motivated by the study of certain problems in computer vision [Kimia *et al.*, 1990; Kimia, 1989]. Indeed, we are interested in studying the problem of shape perception, and explicitly in the reconciliation between the *parts* versus *protrusions* dichotomy in computer vision. This paper is primarily concerned with providing a rigorous basis to these results.

The authors would like to thank Professor Ciprian Foias for some very helpful conversations on this paper, especially about Lemma 4.2.1.

1.1 Notation

We will now set up some of the basic notation and concepts which we will need in what follows.

Let $C(s, t) : S^1 \rightarrow \mathbb{R}^2$ be a family of embedded curves where t denotes time and s parameterizes each curve. We assume that this family evolves according to the evolution equation

$$\frac{\partial C}{\partial t} = \alpha(s, t)\vec{T} + \beta(s, t)\vec{N} \quad (1.1)$$

$$C(s, 0) = C_0(s), \quad (1.2)$$

where \vec{N} is the outward normal, κ is the Gaussian curvature, and α, β are arbitrary functions. For each deformation $\{\alpha, \beta\}$, there exists another deformation $\{0, \beta'\}$ such that the resulting traces of curves are equivalent [Gage, 1986]. Furthermore, we constrain the deformations to be determined by the local geometry of the curve, i.e., β should be a function of curvature [Kimia, 1989]. Therefore, we consider the case $\alpha(s, t) = 0$ where β is typically of the form $\beta(\kappa) = 1 - \epsilon\kappa$. Assume that $C_t = C(., t)$ is a C^2 -classical solution on some interval $[0, t')$ ($t' < \infty$). Thus we are considering C^2 solutions of the system

$$\frac{\partial C}{\partial t} = \beta(\kappa(s, t))\vec{N} \quad (1.3)$$

$$C(s, 0) = C_0(s), \quad (1.4)$$

(Note that we do not rule out the possibility that a C^2 -solution may exist for all $t \geq 0$.)

Let,

$$g(s, t) := \left| \frac{\partial \mathcal{C}}{\partial t} \right| = [x_s^2 + y_s^2]^{1/2},$$

denote the length along the curve. The arc-length parameter \tilde{s} is then defined as

$$\tilde{s}(s, t) := \int_0^s g(\xi, t) d\xi.$$

Let the positive orientation of a curve be defined so that the interior is to the left when traversing the curve. The *tangent*, *curvature*, *normal*, *orientation* and *length* are defined in the standard way. We will take the normal to be pointing outwards, where the inward or outward is determined by the interior, or equivalently by the orientation of the curve. We then have that

$$\begin{aligned} \vec{T} &:= \frac{\partial \mathcal{C}}{\partial \tilde{s}} = \frac{1}{g} \frac{\partial \mathcal{C}}{\partial s}, \\ \kappa &:= \left| \frac{\partial \vec{T}}{\partial \tilde{s}} \right| = \frac{1}{g} \left| \frac{\partial \vec{T}}{\partial s} \right|, \\ \vec{N} &:= \frac{-\frac{\partial \vec{T}}{\partial \tilde{s}}}{\left| \frac{\partial \vec{T}}{\partial \tilde{s}} \right|} = \frac{-1}{\kappa g} \frac{\partial \vec{T}}{\partial s}, \\ \theta &:= \angle(\vec{T}, \vec{x}) \\ L(t) &:= \int_0^{2\pi} g(s, t) ds \end{aligned}$$

We also define a quantity which we will call *length-squared* by

$$L^{(2)}(t) := \int_0^{2\pi} g^2(s, t) ds.$$

Finally, we let

$$\bar{\kappa}(t) := \int_0^{2\pi} |\kappa(s, t)| g(s, t) ds,$$

denote the total absolute Gaussian curvature.

2. On General Curve Flow

In this section we will derive the evolution equation for the tangent \vec{T} , normal \vec{N} , metric g , curvature κ , orientation θ and length L , for families of curves satisfying (1.1).

It is easily established that

$$\begin{cases} \frac{\partial \vec{T}}{\partial s} = -\kappa g \vec{N} \\ \frac{\partial \vec{N}}{\partial s} = \kappa g \vec{T}, \end{cases}$$

which will be used in the following proofs. Note that these evolution equations are for the general deformation $\{\alpha, \beta\}$.

Moreover, we can compute that the metric g evolves as follows:

$$\begin{aligned} \frac{\partial g^2}{\partial t} &= \frac{\partial}{\partial t} \left\langle \frac{\partial \mathcal{C}}{\partial s}, \frac{\partial \mathcal{C}}{\partial s} \right\rangle \\ &= 2 \left\langle \frac{\partial \mathcal{C}}{\partial s}, \frac{\partial}{\partial t} \frac{\partial \mathcal{C}}{\partial s} \right\rangle \\ &= 2 \left\langle \frac{\partial \mathcal{C}}{\partial s}, \frac{\partial}{\partial s} \frac{\partial \mathcal{C}}{\partial t} \right\rangle \\ &= 2 \left\langle g \vec{T}, \frac{\partial}{\partial s} [\alpha \vec{T} + \beta \vec{N}] \right\rangle \\ &= 2 \left\langle g \vec{T}, \frac{\partial \alpha}{\partial s} \vec{T} + \alpha \frac{\partial \vec{T}}{\partial s} + \frac{\partial \beta}{\partial s} \vec{N} + \beta \frac{\partial \vec{N}}{\partial s} \right\rangle \\ &= 2 \left\langle g \vec{T}, \alpha_s \vec{T} - \alpha \kappa g \vec{N} + \beta_s \vec{N} + \beta \kappa g \vec{T} \right\rangle \\ &= 2g[\alpha_s + \beta \kappa g]. \end{aligned}$$

Hence, we see that

$$\frac{\partial g}{\partial t} = \alpha_s + \beta \kappa g.$$

In the special case of $\alpha = 0$,

$$\frac{\partial g}{\partial t} = \beta \kappa g.$$

We will need the following change of partials for computing evolution equations:

$$\begin{aligned} \frac{\partial}{\partial t} \frac{\partial}{\partial \vec{s}} &= \frac{\partial}{\partial t} \left[\frac{1}{g} \frac{\partial}{\partial s} \right] \\ &= \frac{-g_t}{g^2} \frac{\partial}{\partial s} + \frac{1}{g} \frac{\partial}{\partial t} \frac{\partial}{\partial s} \\ &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \frac{\partial}{\partial \vec{s}} + \frac{\partial}{\partial \vec{s}} \frac{\partial}{\partial t}. \end{aligned}$$

Next we have the following evolution equation for tangent:

$$\begin{aligned}
 \frac{\partial \vec{T}}{\partial t} &= \frac{\partial}{\partial t} \frac{\partial \mathcal{C}}{\partial \vec{s}} \\
 &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \frac{\partial \mathcal{C}}{\partial \vec{s}} + \frac{\partial}{\partial \vec{s}} \frac{\partial \mathcal{C}}{\partial t} \\
 &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \vec{T} + \frac{\partial}{\partial \vec{s}} [\alpha \vec{T} + \beta \vec{N}] \\
 &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \vec{T} + \alpha_s \vec{T} + \alpha \frac{\partial \vec{T}}{\partial \vec{s}} + \beta_s \vec{N} + \beta \frac{\partial \vec{N}}{\partial \vec{s}} \\
 &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \vec{T} + \frac{1}{g} \alpha_s \vec{T} - \alpha \kappa \vec{N} + \frac{1}{g} \beta_s \vec{N} + \beta \kappa \vec{T} \\
 &= \frac{1}{g} [\beta_s - \alpha \kappa g] \vec{N}.
 \end{aligned}$$

Similarly, for the normal we see that

$$\begin{aligned}
 \frac{\partial \vec{N}}{\partial t} &= - \left\langle \frac{\partial \vec{N}}{\partial t}, \vec{T} \right\rangle \vec{T} \\
 &= - \left\langle \frac{\partial \vec{T}}{\partial t}, \vec{N} \right\rangle \vec{T} \\
 &= - \frac{1}{g} [\beta_s - \alpha \kappa g] \vec{T}.
 \end{aligned}$$

Next we define the orientation of a curve as the angle the tangent makes with the x -axis. Let $\vec{T} = (\cos(\theta), \sin(\theta))$, so that $\vec{N} = (\sin(\theta), -\cos(\theta))$. Then,

$$\begin{aligned}
 \frac{\partial \vec{T}}{\partial t} &= (-\sin(\theta), \cos(\theta)) \frac{\partial \theta}{\partial t} \\
 &= - \frac{\partial \theta}{\partial t} \vec{N}.
 \end{aligned}$$

Therefore,

$$\begin{cases} \frac{\partial \theta}{\partial t} = \frac{-1}{g} [\beta_s - \alpha \kappa g] \\ \frac{\partial \theta}{\partial s} = \kappa g. \end{cases}$$

As for curvature, we compute that

$$\begin{aligned}
 \frac{\partial \kappa}{\partial t} &= \frac{\partial}{\partial t} \frac{\partial \theta}{\partial \vec{s}} \\
 &= \frac{-1}{g} [\alpha_s + \beta \kappa g] \frac{\partial \theta}{\partial \vec{s}} + \frac{\partial}{\partial \vec{s}} \left[- \frac{\partial \theta}{\partial t} + \alpha \kappa \right]
 \end{aligned}$$

$$\begin{aligned}
&= \frac{-1}{g} [\alpha_s + \beta \kappa g] \kappa - \frac{\partial^2 \beta}{\partial s^2} + \alpha_s \kappa + \alpha \frac{\partial \kappa}{\partial s} \\
&= -\frac{\partial^2 \beta}{\partial s^2} + \alpha \frac{\partial \kappa}{\partial s} - \beta \kappa^2
\end{aligned}$$

For the length, we derive

$$\begin{aligned}
\frac{\partial L}{\partial t} &= \frac{\partial}{\partial t} \int_0^{2\pi} g(s, t) ds \\
&= \int_0^{2\pi} \frac{\partial g(s, t)}{\partial t} ds \\
&= \int_0^{2\pi} [\alpha_s + \beta \kappa g] ds,
\end{aligned}$$

and similarly for length-squared,

$$\begin{aligned}
\frac{\partial L^{(2)}}{\partial t} &= \frac{\partial}{\partial t} \int_0^{2\pi} g^2(s, t) ds \\
&= \int_0^{2\pi} \frac{\partial g^2(s, t)}{\partial t} ds \\
&= \int_0^{2\pi} 2g[\alpha_s + \beta \kappa g] ds.
\end{aligned}$$

We now specialize to the case $\beta(\kappa) = 1 - \epsilon \kappa$ which is a common model frequently used in applications such as flame propagation, crystal growth, among others. First, we can easily show in this case that the *metric* evolves according to

$$\frac{\partial g}{\partial t} = (-\epsilon \kappa^2 + \kappa)g.$$

Second, for the *tangent* and *normal* we have

$$\frac{\partial \vec{T}}{\partial t} = -\frac{\epsilon \kappa_s}{g} \vec{N}.$$

$$\frac{\partial \vec{N}}{\partial t} = \frac{\epsilon \kappa_s}{g} \vec{T}.$$

Next, the *orientation* evolution is governed by

$$\frac{\partial \theta}{\partial t} = \frac{\epsilon \kappa_s}{g}.$$

Similarly, one can show that the evolution equation for *curvature* is

$$\frac{\partial \kappa}{\partial t} = \epsilon \kappa_{ss} + \epsilon \kappa^3 - \kappa^2.$$

Finally, length evolves as

$$L_t = 2\pi - \epsilon \int_0^{2\pi} \kappa^2 g \, ds.$$

It is also useful to further constrain the flow and obtain evolution equations for the particular case of $\epsilon = 0$, or $\beta(\kappa) = 1$, for which we have

$$\begin{aligned} \frac{\partial g}{\partial t} &= \kappa g, \\ \frac{\partial \vec{T}}{\partial t} &= 0, \\ \frac{\partial \vec{N}}{\partial t} &= 0, \\ \frac{\partial \theta}{\partial t} &= 0, \\ \frac{\partial \kappa}{\partial t} &= -\kappa^2, \\ \frac{\partial L}{\partial t} &= 2\pi. \end{aligned}$$

The evolution equation for curvature may be solved explicitly as

$$\kappa(s, t) = \frac{\kappa(s, 0)}{1 + \kappa(s, 0)t}.$$

This implies that the classical solution will fail to exist when

$$t = \frac{-1}{\kappa(s, 0)}.$$

Hence, if the initial curve to (1.3) is convex, the equation will have a classical solution for all time.

The metric equation may also be solved as

$$\begin{aligned} \frac{\partial g}{\partial t} &= \kappa(s, t)g(s, t) \\ \frac{g_t}{g} &= \frac{\kappa(s, 0)}{1 + \kappa(s, 0)t} \\ \frac{\partial \ln(g)}{\partial t} &= \frac{\kappa(s, 0)}{1 + \kappa(s, 0)t} \\ \frac{\partial \ln(g)}{\partial t} &= \frac{\partial \ln(1 + \kappa(s, 0)t)}{\partial t} \\ \ln(g(s, t)) - \ln(g(s, 0)) &= \ln(1 + \kappa(s, 0)t) - \ln(1) \\ \ln(g(s, t)) &= \ln(g(s, 0)) + \ln(1 + \kappa(s, 0)t) \\ g(s, t) &= g(s, 0)(1 + \kappa(s, 0)t) \end{aligned}$$

Hence, the metric changes linearly in time with a curvature dependent coefficient. In particular, for negative curvature the metric decrease to zero at exactly the time the shock is formed.

3. Bounds on Length, Curvature, and Travelled Distance

In this section we give bounds on the length and total absolute curvature for the family defined in Section 1. As before we are particularly interested in the case $\alpha = 0$ and $\beta(s, t) = 1 - \epsilon\kappa(s, t)$ which is of interest in physical applications and recently in computer vision [Kimia, 1989].

3.1 A Bound for Length

LEMMA 3.1: *Let $C(s, t)$ be a solution of (1.3) for $t \in [0, t')$ and $\kappa\beta(\kappa) \leq M$ for all $\kappa \in \mathbb{R}$ (regarding β as a function of κ). Then,*

$$L(t) \leq \min(L(0) + 2\pi t, L(0)e^{Mt}).$$

In particular, for $\beta(\kappa) = 1 - \epsilon\kappa$,

$$L(t) \leq \min(L(0) + 2\pi t, L(0)e^{\frac{t}{\epsilon}}).$$

Proof. We have

$$L_t = 2\pi - \epsilon \int_0^{2\pi} \kappa^2 g \, ds.$$

So,

$$\begin{aligned} L_t &\leq 2\pi, \\ L(t) &\leq 2\pi t + L(0). \end{aligned}$$

Note, the equality holds for $\epsilon = 0$. Alternatively,

$$\frac{\partial L}{\partial t} = \int_0^{2\pi} \beta\kappa g \, ds$$

Since $\kappa\beta(\kappa) \leq M$

$$\begin{aligned} \frac{\partial L}{\partial t} &\leq \int_0^{2\pi} M g \, ds \\ &\leq ML(t) \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{L'(t)}{L(t)} &\leq M \\ (\ln[L(t)])' &\leq M \end{aligned}$$

that is,

$$\ln[L(t)] \leq Mt + \ln(L(0))$$

$$L(t) \leq L(0)e^{Mt}.$$

In particular for $\beta = 1 - \epsilon\kappa$, $M = \frac{1}{4\epsilon}$, and

$$L(t) \leq L(0)e^{\frac{t}{4\epsilon}}.$$

Remarks 1.

(i) In this case, as $\epsilon \rightarrow \infty$, $L(t) \leq L(0)$. It is interesting to observe that the two estimates complement each other: for small ϵ the second estimate is very large making the first estimate more useful. However, for large ϵ , the first estimate is exaggerated making the second estimate more useful.

(ii) The same proof as above shows that the length-squared

$$L^{(2)}(t) := \int_0^{2\pi} g^2 ds$$

is bounded by $L^{(2)}(0)e^{2Mt}$. (This fact could also have been used to bound $L(t)$, but the bound would have been more conservative.)

3.2 A Bound for Total Absolute Curvature

LEMMA 3.2: Let $C(s, t)$ be a solution of (1.3) for $t \in [0, t']$. Suppose that $\kappa\beta(\kappa) \leq M$, and $\beta_\kappa \leq 0$. Then,

$$\bar{\kappa}(t) \leq \bar{\kappa}(0).$$

Proof. Define

$$\hat{q}(t) := \int_0^{2\pi} q(\kappa(s, t))g(s, t) ds,$$

where q is the piecewise smooth convex approximation of $f(x) = |x|$ given by

$$q(x) = \begin{cases} |x| & \text{if } x \geq \frac{1}{n} \\ \frac{1}{2n} + \frac{n}{2}x^2 & \text{if } x \leq \frac{1}{n} \end{cases}$$

Then

$$\begin{aligned} \hat{q}'(t) &= \int_0^{2\pi} q_\kappa(\kappa)\kappa_t g ds + \int_0^{2\pi} q(\kappa)g_t ds \\ &= - \int_0^{2\pi} q_\kappa(\kappa)(\beta_{ss} + \beta\kappa^2)g ds + \int_0^{2\pi} q(\kappa)(\beta\kappa g) ds \\ &= - \int_0^{L(t)} q_\kappa(\kappa)\beta_{ss} d\tilde{s} + \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)](\beta\kappa g) ds. \\ &= -(q_\kappa(\kappa)\beta_s) \Big|_0^{L(t)} + \int_0^{L(t)} q_{\kappa\kappa}(\kappa)\kappa_s \beta_s d\tilde{s} + \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)](\beta\kappa g) ds. \end{aligned}$$

$$\begin{aligned}
&= \int_0^{L(t)} q_{\kappa\kappa}(\kappa) \kappa_s \beta_s d\tilde{s} + \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)] (\beta \kappa g) ds. \\
&= \int_0^{L(t)} q_{\kappa\kappa}(\kappa) \kappa_s^2 \beta_\kappa d\tilde{s} + \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)] (\beta \kappa g) ds.
\end{aligned}$$

Since $\beta_\kappa \leq 0$ and convexity of q requires $q'' \geq 0$, we have

$$\tilde{q}'(t) \leq \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)] (\beta \kappa g) ds.$$

Note that

$$\bar{\kappa}(t) \leq \hat{q}(t),$$

so that a bound on \hat{q} is a bound on $\bar{\kappa}$. Moreover, we have that

$$0 \leq q(x) - xq'(x) \leq \begin{cases} 0 & \text{if } x \geq \frac{1}{n} \\ \frac{1}{2n} & \text{if } x \leq \frac{1}{n} \end{cases}$$

Now since

$$\beta(\kappa) \kappa \leq M,$$

and $[q(\kappa) - \kappa q_\kappa(\kappa)] \geq 0$,

$$\begin{aligned}
\tilde{q}'(t) &\leq M \int_0^{2\pi} [q(\kappa) - \kappa q_\kappa(\kappa)] g ds. \\
&\leq M \frac{1}{2n} \int_0^{2\pi} g ds. \\
&\leq \frac{M}{2n} L(t),
\end{aligned}$$

and so,

$$\bar{\kappa}'(t) \leq \frac{M}{2n} L(t)$$

Since by lemma 3.1 the length is finite, letting $n \rightarrow \infty$, we get

$$\bar{\kappa}'(t) \leq 0,$$

or

$$\bar{\kappa}(t) \leq \bar{\kappa}(0).$$

■

Remark 2. The conditions of the above lemma holds for $\beta(\kappa(s, t)) = 1 - \epsilon \kappa(s, t)$, for $\epsilon > 0$.

LEMMA 3.3: *Let a family of curves satisfy (1.3) with convex initial condition, i.e. $\kappa(s, 0) \geq 0$. Then, $\bar{\kappa}(t) = \bar{\kappa}(0) = 2\pi$ and the curve remains convex for all times.*

Proof. For convex curves,

$$\bar{\kappa}(s, 0) = \int_0^{2\pi} \kappa g \, ds = 2\pi.$$

There is some neighborhood of time such that $\kappa(s, t) > 0$ for all s . Therefore,

$$\begin{aligned} \frac{\partial \bar{\kappa}}{\partial t} &= \frac{\partial}{\partial t} \int_0^{2\pi} \kappa g \, ds \\ &= \int_0^{2\pi} [\kappa_t g + \kappa g_t] \, ds \\ &= \int_0^{2\pi} [(-\beta_{ss} - \beta \kappa^2)g + \kappa \beta \kappa g] \, ds \\ &= - \int_0^{2\pi} \beta_{ss} g \, ds \\ &= - \int_0^{L(t)} \beta_{ss} \, d\tilde{s} \\ &= -\beta_s \Big|_0^{L(t)} \\ &= 0. \end{aligned}$$

Hence,

$$\bar{\kappa}(t) = 2\pi.$$

Since only convex curves can satisfy this condition, the evolved curve must be convex. ■

LEMMA 3.4: *Let a family of curves satisfy (1.3) for which $\beta_\kappa < 0$. Then, if $\kappa_{\tilde{s}}(\tilde{s}, t) \neq 0$ for all \tilde{s} and $0 \leq t < t'$*

$$\bar{\kappa}(t) < \bar{\kappa}(0)$$

. For $\beta_\kappa = 0$,

$$\bar{\kappa}(t) = \bar{\kappa}(0).$$

Proof. This proof is due to Sethian [Sethian, 1985] and we include it here for completeness with a few changes.

Without loss of generality pick the starting point $\tilde{s} = 0$ to be a zero of curvature such that positive curvature begins in the positive direction of the curve. Partition the interval $[0, L(t)]$ into $n + 1$ maximal subintervals,

$$\{[\tilde{s}_0 = 0, \tilde{s}_1], [\tilde{s}_1, \tilde{s}_2], \dots, [\tilde{s}_i, \tilde{s}_{i+1}], \dots, [\tilde{s}_n, \tilde{s}_{n+1} = L(t)]\},$$

such that $\kappa(\tilde{s})$ is entirely positive, negative, or zero in the interval $(\tilde{s}_i, \tilde{s}_{i+1})$ for $i = 1, 2, \dots, n$. Then, \tilde{s}_i are zeros of curvature, $\kappa(\tilde{s}_i) = 0$. Note further that in general $s_i = s_i(t)$. Also, let

$$\rho(\kappa, [a, b]) := \begin{cases} 1 & \text{if } \kappa(\tilde{s}) > 0 \text{ for } \tilde{s} \in (a, b) \\ 0 & \text{if } \kappa(\tilde{s}) = 0 \text{ for } \tilde{s} \in (a, b) \\ -1 & \text{if } \kappa(\tilde{s}) < 0 \text{ for } \tilde{s} \in (a, b) \end{cases}$$

Then,

$$\begin{aligned}
 \frac{\partial \bar{\kappa}}{\partial t} &= \frac{\partial}{\partial t} \int_0^{2\pi} |\kappa| g \, ds \\
 &= \frac{\partial}{\partial t} \sum_{i=0}^n \int_{s_i}^{s_{i+1}} |\kappa| g \, ds \\
 &= \sum_{i=0}^n \int_{s_i}^{s_{i+1}} \frac{\partial[|\kappa|g]}{\partial t} \, ds + \sum_{i=0}^n |\kappa(s_{i+1}, t)| g(s_{i+1}, t) \frac{\partial s_{i+1}}{\partial t} - \sum_{i=0}^n |\kappa(s_i, t)| g(s_i, t) \frac{\partial s_i}{\partial t} \\
 &= \sum_{i=0}^n \int_{s_i}^{s_{i+1}} [|\kappa|_t g + |\kappa| g_t] \, ds
 \end{aligned}$$

Since intervals for which curvature is uniformly zero do not contribute to the sum, we will discount them. Without loss of generality assume otherwise in the following.

$$\begin{aligned}
 \frac{\partial \bar{\kappa}}{\partial t} &= \sum_{i=0}^n \int_{s_i}^{s_{i+1}} \left[\frac{\kappa \kappa_t}{|\kappa|} g + |\kappa| \beta \kappa g \right] \, ds \\
 &= \sum_{i=0}^n \int_{s_i}^{s_{i+1}} \left[\frac{\kappa}{|\kappa|} (-\beta_{ss} - \beta \kappa^2) + |\kappa| \beta \kappa \right] \, d\tilde{s} \\
 &= \sum_{i=0}^n \int_{s_i}^{s_{i+1}} \left[\frac{\kappa}{|\kappa|} (-\beta_{ss}) \right] \, d\tilde{s} + \sum_{i=0}^n \int_{s_i}^{s_{i+1}} \left[\frac{-\beta \kappa^3}{|\kappa|} + \beta \kappa |\kappa| \right] \, d\tilde{s} \\
 &= - \sum_{i=0}^n \int_{s_i}^{s_{i+1}} [\rho(\kappa, [\tilde{s}_i, \tilde{s}_{i+1}])] \beta_{ss} \, d\tilde{s} \\
 &= - \sum_{i=0}^n \rho(\kappa, [\tilde{s}_i, \tilde{s}_{i+1}]) \beta_s \Big|_{s_i}^{s_{i+1}} \\
 &= - \sum_{i=0}^n \rho(\kappa, [\tilde{s}_i, \tilde{s}_{i+1}]) [\beta_s(\tilde{s}_{i+1}) - \beta_s(\tilde{s}_i)]
 \end{aligned}$$

Now, by our original assumption $\rho(\kappa, [\tilde{s}_0, \tilde{s}_1]) = 1$. The next interval, then, has negative curvature and $\rho(\kappa, [\tilde{s}_1, \tilde{s}_2]) = -1$. Since zeros of curvature must pair up, $\rho(\kappa, [\tilde{s}_n, \tilde{s}_{n+1}]) = -1$. In short, $\rho(\kappa, [\tilde{s}_i, \tilde{s}_{i+1}]) = (-1)^i$. Therefore,

$$\begin{aligned}
 \frac{\partial \bar{\kappa}}{\partial t} &= - \sum_{i=0}^n (-1)^i [\beta_s(\tilde{s}_{i+1}) - \beta_s(\tilde{s}_i)] \\
 &= - \sum_{i=0}^n (-1)^i \beta_s(\tilde{s}_{i+1}) + \sum_{i=0}^n (-1)^i \beta_s(\tilde{s}_i) \\
 &= \sum_{j=1}^{n+1} (-1)^j \beta_s(\tilde{s}_j) + \sum_{i=0}^n (-1)^i \beta_s(\tilde{s}_i) \\
 &= \sum_{j=0}^n (-1)^j \beta_s(\tilde{s}_j) + \sum_{i=0}^n (-1)^i \beta_s(\tilde{s}_i) \\
 &= 2 \sum_{i=0}^n (-1)^i \beta_s(\tilde{s}_i)
 \end{aligned}$$

$$= 2 \sum_{i=0}^n (-1)^i \beta_{\kappa}(\tilde{s}_i) \kappa_{\beta}(\tilde{s}_i)$$

Now, since $\kappa_{\beta}(\tilde{s}_i)$ has sign $(-1)^i$ if $\beta_{\kappa} < 0$, then

$$\frac{\partial \bar{\kappa}}{\partial t} < 0$$

so that

$$\bar{\kappa}(t) < \bar{\kappa}(0).$$

However, if $\beta_{\kappa} = 0$, such as the case with $\beta = 1$,

$$\bar{\kappa}(t) = \bar{\kappa}(0).$$

■

Remark 3.

In conclusion, convex curves remain convex and $\bar{\kappa}(t) = 2\pi$ for all deformations. For nonconvex curves and $\beta = 1$, we have $\bar{\kappa}(t) = \bar{\kappa}(0)$. Note therefore that for all curves, the deformation $\beta = 1$ does not alter the total absolute curvature. Finally, for nonconvex curves and deformations for which $\beta_{\kappa} < 0$, such as $\beta = 1 - \epsilon\kappa$ with $\epsilon > 0$, we have

$$\bar{\kappa}(t) < \bar{\kappa}(0).$$

This describes the important role of ϵ in the deformation as one of reducing the total absolute curvature. Note that for $\beta = -\epsilon\kappa$, the deformation will evolve an embedded curve to a circle [Gage and Hamilton, 1986; Grayson, 1987].

4. How Far Can a Curve Travel?

In this section we will address the key issue of how far the evolved curve can be away from the initial curve. First we derive a relationship between the distance of a point from a curve and the curvature of the curve at the nearest point. Second, the rate of change of distance of a point to a curve is related to the speed of the curve at its nearest point. This result holds for nonshock points. Third, the distance of a point from a curve bounds the rate of change of that distance with time. Fourth, we show that a curve can not travel too far pointwise. Fifth, for any time neighbourhood for which the curve does not travel beyond ϵ , we constrain its expansion as a function of time. From these we conclude that two curves close in time are close in their Hausdorff distance. Finally, a theorem shows that the limit of curvature evolution exists, and using the above we can even bound the total Gaussian curvature.

4.1 On the Distance Travelled

In what follows we will limit ourselves to the case

$$\beta(\kappa) = 1 - \epsilon\kappa.$$

For a subset $S \subset \mathbb{R}^2$, let $N_\delta(S)$ denote a closed δ -neighborhood. Define the *signed distance* of a point from a curve (regarded as a point set in \mathbb{R}^2) as

$$d(p, C_t) := \begin{cases} \inf \{d(p, q) | q \in C_t\} & \text{if } p \text{ is outside } C_t \\ -\inf \{d(p, q) | q \in C_t\} & \text{otherwise,} \end{cases}$$

where *outside* is the region to the right of the curve as one traverses the curve in the positive orientation. In this section we will consider an arbitrary point in the plane, p , and consider its relation to the curve C_t . Set

$$d(t) := d(p, C_t)$$

LEMMA 4.1: *Let $p \notin C$ be a point in \mathbb{R}^2 . Let q be the closest point on the curve to p . (Note that q exists by compactness.) Then,*

$$\begin{cases} \kappa(q) \geq \frac{-1}{d(p, q)} & \text{if } p \text{ is outside } C \\ \kappa(q) \leq \frac{-1}{d(p, q)} & \text{if } p \text{ is inside } C. \end{cases}$$

Proof. Set $d = d(p, q)$. First suppose the point p is outside the curve, so that $d > 0$. Let q be the closest point on the curve C to p . Consider the circle of radius d and center p which is tangent to the curve at q . We have two separate cases: (i) the curve has non-negative curvature at q . In this case, $\kappa(q) \geq \frac{-1}{d}$, trivially. (ii) the curve has negative curvature at q . In this case, the curve C lies entirely outside the circle. In order to see this suppose to the contrary; then there exist points on C closer to p than d , a contradiction. Therefore, the curvature of the circle $\frac{1}{d}$ is

greater than the curvature of the curve at q , i.e. $\frac{1}{d} \geq -\kappa(q)$. Therefore, for both cases when the point p is outside the curve C , we have

$$\kappa(q) \geq \frac{-1}{d}$$

Now, as required suppose the point p is inside the curve C , $d < 0$, and q the closest point of the curve C to it. Again, consider the circle of radius $-d$ and center p which is tangent to the curve at q . Once more we have two cases: (i) the curve has non-positive curvature at q . Trivially, then $\kappa(q) \leq \frac{-1}{d}$. (ii) the curve has positive curvature at q . In this case, the circle again lies entirely within the curve, touching it only at q , so that, the curvature of the circle $\frac{-1}{d}$ is greater than the curvature of the curve $\kappa(q)$. Then,

$$\kappa(q) \leq \frac{-1}{d}.$$

LEMMA 4.2: Let p be a point in \mathbb{R}^2 where C_t deforms along the normal according to (1.3). If $\frac{\partial \theta}{\partial t}$ can be bounded, then

$$d'(t) = -\frac{\partial C}{\partial t}.$$

Proof. Let $q(s, t)$ be the closest point to p on C_t and $q(s + \delta s, t + \delta t)$ the closest point on $C_{t+\delta t}$ to p . Since the line $(p, q(s, t))$ is normal to C_t , the point $q(s, t + \delta t)$ is on this line a distance $\frac{\partial C}{\partial t} \delta t$ from $q(s, t)$. Consider the triangle with vertices $p, q(s, t + \delta t), q(s + \delta s, t + \delta t)$, where the angle at p is denoted by $\delta \theta$. Then,

$$\begin{aligned} \frac{d(t + \delta t) - d(t)}{\delta t} &= \frac{1}{\delta t} \left[\frac{(d(t) - \frac{\partial C}{\partial t} \delta t)}{\cos(\delta \theta)} - d(t) \right] \\ &= \frac{1}{\cos(\delta \theta) \delta t} [2d(t) \sin^2(\delta \theta / 2) - \frac{\partial C}{\partial t} \delta t] \\ &= -2d(t) \frac{\sin^2(\delta \theta / 2)}{\cos(\delta \theta) \delta t} - \frac{\frac{\partial C}{\partial t}}{\cos(\delta \theta)}. \end{aligned}$$

In the limit, $\delta t \rightarrow 0$, $\delta \theta \rightarrow 0$. From the equations for rate of change of orientation with \sin Section 2,

$$\frac{\partial \theta}{\partial t} = \frac{-\beta_s}{g}.$$

Since θ_t can be bounded

$$d'(t) = -\frac{\partial C}{\partial t}.$$

Remark 4. Observe that as t approaches the time of shock formation, $g(s, t)$ goes to zero, and therefore the condition of Lemma 4.2 do not hold in the limit. -

LEMMA 4.3: Let C_t be a solution of (1.3) where $\beta(\kappa) = 1 - \epsilon \kappa$. Let p be a point in \mathbb{R}^2 . If $d(t) = d(p, C_t) \leq \epsilon$, then,

$$d'(t)d(t) \geq -2\epsilon.$$

Proof. Let $q(t)$ be the closest point on the curve C_t to p . First, consider the case where the point p is outside the curve C_t . Then, by Lemma 4.1, $\kappa(q(t)) \geq \frac{-1}{d(t)}$. Consequently,

$$\begin{aligned} d'(t) &= \epsilon \kappa - 1 \\ &\geq \frac{-\epsilon}{d(t)} - 1 \\ &\geq \frac{-2\epsilon}{d(t)}, \end{aligned}$$

since $d(t)$ is positive and $d(t) \leq \epsilon$. Hence we can conclude that

$$d'(t)d(t) \geq -2\epsilon.$$

Now, consider the case where the point p is inside the curve which implies $\kappa(q(t)) \leq \frac{-1}{d(t)}$. Then,

$$\begin{aligned} d'(t) &= \epsilon \kappa - 1 \\ &\leq \frac{-\epsilon}{d(t)} - 1 \\ &\leq \frac{-2\epsilon}{d(t)}, \end{aligned}$$

since $d(t)$ is negative and $d(t) \leq \epsilon$. Once again, we can conclude that

$$d'(t)d(t) \geq -2\epsilon.$$

■

LEMMA 4.4: Consider a curve C_0 evolving through a function of curvature as in (1.3), with $\beta(\kappa) = 1 - \epsilon\kappa$. Then for each ϵ , there exists $\bar{t} = \bar{t}(\epsilon, C_0) > 0$, such that for each $p \in C_{\hat{t}}$ with $0 \leq \hat{t} < \bar{t}$ we have that

$$d(t) \leq \epsilon \quad \forall 0 \leq t < \bar{t}.$$

Proof. Let $p \in C_{\hat{t}}$ for some $0 \leq \hat{t} < \bar{t}$. Note that $1 - \epsilon\kappa(q(t), t)$ is the speed of the point $q(t)$ on C_t in the evolving family. Now since $\kappa(s, t)$ is a periodic solution of a polynomial reaction-diffusion equation with analytic coefficients and smooth initial condition, there exists an interval $[0, t_1]$ such that $\kappa(s, t)$ is uniformly bounded as a function in t say by M . Therefore on $[0, t_1]$

$$|1 - \epsilon\kappa(q(t), t)| \leq 1 + \epsilon M.$$

Thus any point $q(0) \in C_0$ cannot have travelled more than distance ϵ from C_0 in time

$$t_2 := \frac{\epsilon}{1 + \epsilon M}.$$

Set

$$\bar{t} := \min(t_1, t_2).$$

Since $p \in C_t$ for some $0 \leq \hat{t} < \bar{t}$, the lemma is proven. ■

THEOREM 4.5: Consider a curve C_0 evolving through a function of curvature as in 1.3 with $\beta(\kappa) = 1 - \epsilon\kappa$. Let $\bar{t} = \bar{t}(\epsilon, C_0)$ be as in the previous lemma. Then

$$C_t \subset N_{\sqrt{4\epsilon t}}(C_0),$$

for all $t \in [0, \bar{t}]$.

Proof. From Lemma 4.4, given ϵ there exists \bar{t} such that for all $t \in [0, \bar{t}]$ and for all $p \in C_t$ we have

$$d(t) = d(p(t), C_t) \leq \epsilon.$$

Now, by Lemma 4.3,

$$\left[\frac{d^2(t)}{2}\right]' = d'(t)d(t) \geq 2\epsilon,$$

for all $t \in [0, \bar{t}]$. By integration

$$d(t) \leq \sqrt{4\epsilon t}.$$

Hence, we have shown there is \bar{t} such that for $t \in [0, \bar{t}]$ we have

$$C_t \subset N_{\sqrt{4\epsilon t}}(C_0).$$

■

Remark 5. For d_H the Hausdorff metric defined on compact subsets of \mathbb{R}^2 , from the above theorem we have that

$$d_H(C_t, C_0) \leq \sqrt{4\epsilon t}.$$

4.2 Limits of Classical Solutions

LEMMA 4.6: Let $C_t : S^1 \rightarrow \mathbb{R}^2$ be a family of C^2 functions with uniformly bounded length-squared $L^{(2)}(t)$. Then, C_t is uniformly equicontinuous.

Proof. The lemma is a simple modification of the result for functions with bounded derivatives. Note that

$$C(s, t) - C(s_0, t) = \int_{s_0}^s \frac{\partial C}{\partial \sigma} d\sigma.$$

Therefore if L is the uniform bound on the length-squared

$$\begin{aligned} \|C(s, t) - C(s_0, t)\| &\leq \\ \int_{s_0}^s \left\| \frac{\partial C}{\partial \sigma} \right\| d\sigma &= \int_{s_0}^s \left[\left(\frac{\partial x}{\partial \sigma} \right)^2 + \left(\frac{\partial y}{\partial \sigma} \right)^2 \right]^{1/2} d\sigma \end{aligned}$$

$$\leq \left[\int_{s_0}^s \left(\frac{\partial x}{\partial \sigma} \right)^2 + \left(\frac{\partial y}{\partial \sigma} \right)^2 d\sigma \right]^{1/2} \leq L|s - s_0|^{1/2},$$

that is, the family is equicontinuous with Hölder constant L and exponent $1/2$. ■

THEOREM 4.7: Consider a curve C_0 evolving through a function of curvature as in (1.3). Then,

$$\lim_{t \rightarrow t'} C_t = C^*,$$

in the Hausdorff metric. The curve C^* regarded as a mapping $C^* : S^1 \rightarrow \mathbb{R}^2$ is Hölder continuous with exponent $1/2$.

Proof. Since the lengths-squared of the curves C_t are uniformly bounded (see Lemma 3.1 and Remark 1(ii)) regarding each $C_t : S^1 \rightarrow \mathbb{R}^2$, we can apply the Lemma 4.6 to the family

$$\{C_t\}_{t \in [0, t']}$$

to conclude that it is equicontinuous. Moreover from Theorem 4.5, the curves lie in a compact region. Thus by the Arzela-Ascoli theorem and the proof of Lemma 4.6, there exists a uniformly convergent subsequence $C_{t_n} \rightarrow C^*$, where $C^* : S^1 \rightarrow \mathbb{R}^2$ is a Hölder continuous function with exponent $1/2$. (The Hölder continuity of the limit follows from the fact that the family is equicontinuous and Hölder continuous. C^* will also denote the corresponding curve.) Thus as compact subsets of the plane, we have that $C_{t_n} \rightarrow C^*$ in the Hausdorff metric.

To complete the proof, we need to show that all the $C_t \rightarrow C^*$ (in the Hausdorff metric) as $t \rightarrow t'$. Let $\delta > 0$, and choose t_n such that

$$C_{t_n} \subset N_{\delta/2}(C^*)$$

and

$$t_n > t' - \frac{\delta^2}{16\epsilon}.$$

(We choose $\delta > 0$ sufficiently small so that $t' - \frac{\delta^2}{16\epsilon} > 0$.) Note that for all $t \in [t_n, t']$, we have

$$t - t_n < \frac{\delta^2}{16\epsilon}.$$

Therefore, by Theorem 4.5

$$\begin{aligned} C_t &\subset N_{\sqrt{4 \cdot \frac{\delta^2}{16\epsilon}}}(C_{t_n}), \\ C_t &\subset N_{\delta/2}(C_{t_n}), \\ C_t &\subset N_{\delta}(C^*), \end{aligned}$$

as required. ■

Remarks 5.

(i) Note that since C^* is Hölder continuous, and since the total Gaussian curvature of the family is uniformly bounded (see Lemma 3.2.1), C^* will have finite total Gaussian curvature.

(ii) From the above results, we have a fairly complete picture about the the classical evolution of a family of curves with

$$\beta(\kappa) = 1 - \epsilon\kappa \quad \epsilon \geq 0.$$

In a future paper, we would like to study weak solutions of this family (after possible shock formation) with an emphasis on the physical applications to vision.

References

- [S. Angenent, 1989] S. Angenent, "Parabolic Equations for Curves and Surfaces I, II," University of Wisconsin-Madison Technical Summary Reports, #89-19, 89-24, 1989.
- [Ben-Jacobi et al., 1983] E. Ben-Jacobi, N. Goldenfeld, J. Langer, and G. Schon, "Dynamics of interfacial pattern formation," *Physical Review Letters*, vol. 51, no. 21, pp. 1930-1932, 1983.
- [Brower et al., 1984] R. C. Brower, D. A. Kessler, J. Koplik, and H. Levine, "Geometrical models of interface evolution," *Physical Review A*, vol. 29, pp. 1335-1342, 1984.
- [Gage and Hamilton, 1986] M. Gage and R. S. Hamilton, "The heat equation shrinking plane curves," *J. Differential Geometry*, vol. 23, pp.69-96, 1986.
- [Gage, 1986] M. Gage, "On an area-preserving evolution equation for plane curves," *Contemp. Math.*, vol. 51, pp. 51-62, 1986.
- [Grayson, 1987] M. A. Grayson, "The heat equation shrinks embedded plane curves to round points," *J. Differential Geometry*, vol. 26, pp. 285-314, 1987.
- [Kimia et al., 1990] B. B. Kimia, A. Tannenbaum, and S. W. Zucker, "Toward a computational theory of shape: An overview," in *Lecture Notes in Computer Science*, vol. 427, pp. 402-407, 1990.
- [Kimia, 1989] B. B. Kimia, "Conservations laws and a theory of shape," Ph.D. dissertation, McGill Centre for Intelligent Machines, McGill University, Montreal, Canada, 1989.
- [Langer, 1980] J. S. Langer, "Instabilities and pattern formation in crystal growth," *Reviews of Modern Physics*, vol. 52, no. 1, pp. 1-28, 1980.
- [Lax, 1971] P. D. Lax, "Shock waves and entropy," in *Contributions to Nonlinear Functional Analysis*, edited by E. Zarontello, pp. 603-634, Academic Press, New York, 1971.
- [Lax, 1973] P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, SIAM Regional Conference Series in Applied Mathematics, Philadelphia, 1973.
- [Osher and Sethian, 1988] S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12-49, 1988.
- [Sethian, 1985] J. A. Sethian, "Curvature and the evolution of fronts," *Comm. Math. Physics*, vol. 101, pp. 487-499, 1985.
- [Sivashinsky, 1977] G. I. Sivashinsky, "Nonlinear analysis of hydrodynamic instability in laminar flames, I: Derivation of basic equations," *Acta Astronautica*, vol. 4, pp. 1177-1206, 1977.

Appendix VII

David, C., and Zucker, S.W., Potentials, valleys, and dynamic global coverings, *Int. J. Computer Vision*, in press.

Potentials, Valleys, and Dynamic Global Coverings¹

Chantal David & Steven W. Zucker²

McGill University
Research Centre for Intelligent Machines
and
Department of Electrical Engineering
3480 University St.
Montréal, Québec, Canada

Abstract

We present a new approach to effect the transition between local and global representations. It is based on the notion of a covering, or a collection of objects whose union is equivalent to the full one. The mathematics of computing global coverings are developed in the context of curve detection, where an intermediate representation (the tangent field) provides a reliable local description of curve structure. This local information is put together globally in the form of a potential distribution. The elements of the covering are then short curves, each of which evolves in parallel to seek the valleys of the potential distribution. The initial curve positions are also derived from the tangent field, and their evolution is governed by variational principles. When stationary configurations are achieved, the global dynamic covering is defined by the union of the local dynamic curves.

1. Introduction

One of the key questions in computational vision is how to effect the transition from local representations to global ones. For example, edge operators give an indication of local edge position and orientation; how can they be threaded together into a global contour? A popular solution is commonly called "edge following" or contour tracing [Ballard and Brown, 1982; Levine, 1985] and it works as follows: from a starting edge position, move in the direction indicated by the orientation until the next edge point is encountered, then move in that direction, and so on until the final point is reached. But such algorithms are inherently sequential, and result in myriad "garden paths" if any errors or noise are present. Global parameters, such as starting and ending points; total number of edge points; or total contour length, aid in their application, but are rarely available (except through user interaction, e.g., Kass, Terzopoulos, and Witkin, [1987]). To make matters worse, inferring global parameters, such as the total length of a curve, raises problems as difficult as global

¹ Research supported by the Natural Sciences and Engineering Research Council of Canada and the Air Force Office of Scientific Research

² Fellow, Canadian Institute for Advanced Research

curve detection itself! Other methods, more sophisticated than simple edge following, been developed in an attempt to apply global constraints, and include minimizing properties such as contrast variation along the curve, total curvature [Montanari, 1971; Martelli, 1974], smoothness across scale space [Lowe, 1988], energy within plate and membrane models [Blake and Zisserman, 1987; Terzopoulos, 1986], or applying top down information [Tsotsos, 1987; Draper *et al.*, 1989]. However, few of the objects in our visual world are made from bent plates, and generally problems arise in verifying when such constraints are valid globally. It is almost as if the cart is being put before the horse, in that (hypothetical) global information is being used to force local decisions.

In an attempt to turn this situation around, we introduce a new class of algorithms for synthesizing contours from local representations of their differential structure. These algorithms are inherently parallel and are natural for vision applications. The mathematical idea behind these new algorithms is to compute properties of global structures indirectly by computing a *covering* of them. Loosely speaking, a covering of an object is a collection of different objects whose union is equivalent to it. In our case, we shall define a covering of a global curve that consists in a collection of short, overlapping curves. Each of these derives from a local representation of image information (which we call a tangent field) and then evolves like the "snakes" of Kass *et. al* [1988] according to a potential distribution also derived from the tangent field. But unlike standard "snakes", ours are local objects; it is their union that comprises the cover. Each element of the cover evolves dynamically and in parallel, and no external global information is required. It is all available by construction in the potential distribution. For these reasons we refer to our representation of global curves as *dynamic coverings*.

The idea of computing dynamic coverings developed in a biologically-oriented project of curve detection, a brief overview of which is provided next (see also [Zucker, Dobbins and Iverson, 1989]). Within this context, the key to computing dynamic coverings is to impose two representations between the image and global curves: a tangent field and a potential distribution. The tangent field (defined in Sec. 2) represents the local properties reliably, and the potential distribution puts them together globally. These intermediate representations thus break the enormous gulf between images and global contours into more manageable "chunks", and represent another key difference between our approach and other recent ones [Kass *et al*, 1987; Blake and Zisserman, 1987]. We infer the tangent field from the image, and then build the potential distribution and snakes from it; other researchers attempt to build the potential distribution directly from the image. We can actually calculate the generators for the potential distribution from the equivalence class of curves that map into each tangent field entry.

The paper is organized as follows. Following the overview of our approach to curve detection (Sec. 2), we develop the preliminary notions of potential distributions (Sec. 3) and deformable curves (Sec. 4). We then get to the heart of the paper in Sec. 5, where the precise definition of a global covering is developed. Examples of the dynamic computation of global coverings are finally shown in Sec. 6, for both artificial and natural images.

2. Overview of the algorithm

The algorithm for curve detection consists of two distinct stages, the first leading to a coarse description of the local structure of curves, and the second to a much finer, global one [Zucker, David, Dobbins, and Iverson, 1988].

2.1 Stage 1: Inferring the tangent field

Orientation selection is the inference of a local description of a curve everywhere along it. Formally, this amounts to inferring the trace of the curve, or the set of points (in the image) through which the curve passes, its quantized tangent and curvature at those points, and the discontinuities [Zucker, 1985]. Through all this work, we will refer to such information as the *tangent field*.

This first stage of orientation selection is in turn modeled as a two step process:

Step 1.1. *Initial Measurement* of the local fit at each point to a model of orientation and curvature. The form of the initial measurements is biologically motivated, and a model of endstopped simple cells, which encode both the orientation and curvature estimates [Dobbins, Zucker, and Cynader, 1987] is used to perform them. The local measurements are quantized into discrete classes of orientation and curvature. However, since the local measurements are inherently inaccurate, we require

Step 1.2. *Interpretation* of the local measurements such that they become globally consistent. Curvature, which relates neighboring tangents, is used to define a natural functional to be minimized in order to attain consistency [Parent and Zucker, 1989; Iverson and Zucker, 1988]. Relaxation labelling [Hummel and Zucker, 1983] provides the formal framework for this.

A complete treatment of this first stage can be found in Iverson [1988].

At the end of the first stage, the tangent field gives the discrete trace of the curves in an image. The two steps above are necessary to guarantee that it contains reliable information. More precisely, the *tangent field* is a set of n 3-tuples,

$$\{((x_i, y_i), \theta_i, \kappa_i) \mid (x_i, y_i) \in I, \theta_i \in \Theta, \kappa_i \in \mathcal{K}\}$$

$i = 1, \dots, n$, where n is the number of trace points in the image I , (x_i, y_i) denotes the (quantized) grid coordinate of the i^{th} trace point indexed over the image I , and θ_i and κ_i are its quantized orientation and curvature, respectively. Typically, Θ consists of eight orientation classes, and \mathcal{K} of five curvature classes (see Fig. 1(a)). Thus the tangent field

summarizes the local differential properties of curves in what Zucker [1987] described as a qualitative manner.

Two topological properties of the tangent field are important: (1) each curve in the image is represented as a connected (in the discrete sense) set of tangent field entries; and (2) discontinuities, intersections and bifurcations are represented as multiple tangent field entries at identical coordinates. To elaborate on (1), recall that, except on the border, each point (x, y) of a (digital) image has four horizontal and vertical neighbors, and four diagonal neighbors; these eight points are called the (3×3) *neighborhood* of (x, y) . A *chain* between two points (x_1, y_1) and (x_2, y_2) is a sequence of points where the first one is (x_1, y_1) , the last one is (x_2, y_2) , and each point of the sequence is in the neighborhood of the preceding one. A set of points of an image is then said to be connected when there exists a chain between any two points of the set [Rosenfeld and Kak, 1982]. For an elaboration of (2), see [Zucker et al., 1988].

2.2 Stage 2: Inferring a covering of the curve

Since the tangent is the first derivative of a curve (with respect to arc length), the global curve can be recovered as an integral through the tangent field. Such a view typically leads to sequential recovery (numerical integration) algorithms (e.g., contour following). But these algorithms require starting points and some amount of topological structure (i.e., which tangent point follows which), and there are immense difficulties inherent in obtaining these parameters directly from the image. Moreover, because we are interested in biologically plausible models, we shall require that all the algorithms be parallel.

The key idea behind our approach is to recover the global curve by computing a *covering* that is sufficient for obtaining any global properties of the curve required for subsequent visual processing. The elements of the covering are short curves, initially born at each (discrete) tangent location with unit length, but which then evolve according to a potential distribution constructed from the tangent field. Thus the algorithm that we are proposing can be viewed, mathematically, as approximating the computation of integral curves through a direction field in a parallel manner with very non-standard initial conditions (provided, of course, that multiple entries at the same position are properly interpreted).

Again, there are two conceptually distinct steps to Stage 2 of the algorithm:

Step 2.1. *Constructing the Potential Distribution* from the discrete tangent field. Assuming that each tangent in the tangent field gives a coarse hypothesis about the behavior of the curve in its neighborhood, for each tangent field entry we represent this class of hypotheses by an elongated (in the tangent direction) 2D-Gaussian. This weighting function indicates the (equivalence) class of curves which project to each tangent field entry in extending the discrete tangent in both length and width. The potential distribution is the pointwise summation of these n (one per tangent field entry)

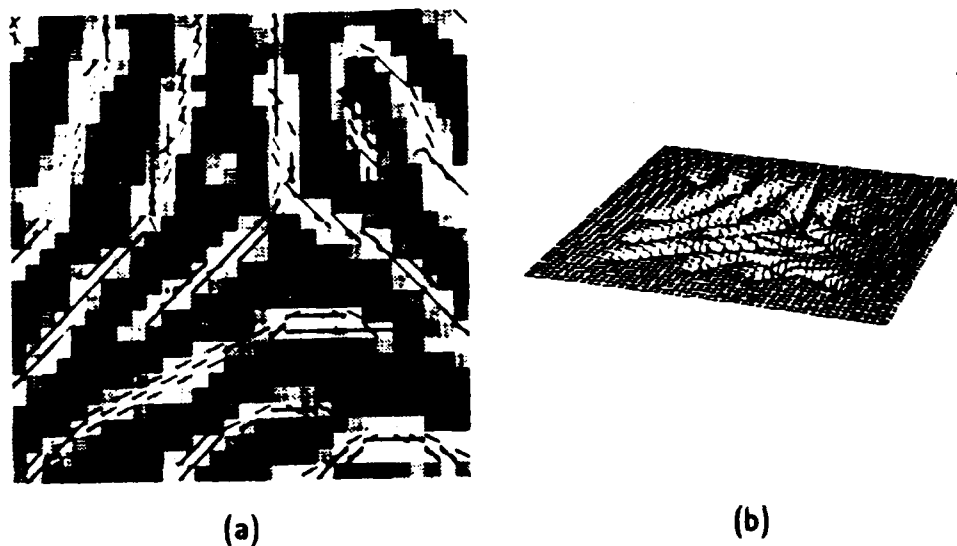


Figure 1 Illustration of the two intermediate structures between the image and the global curves: the tangent field and the potential distribution. The image is a section of a fingerprint image, where each individual pixel is shown as a little block. (a) The tangent field, superimposed onto the original image. Each tangent is represented as a unit length straight line in the tangent direction (curvature is not displayed). (b) The potential distribution, computed from the discrete tangent field by pointwise summation of elongated 2D Gaussians. Notice the presence of smooth valleys, corresponding to the curves in the image.

Gaussians. The key feature of the potential distribution is the presence of smooth valleys, constructed by summation of nearby Gaussians. These valleys indicate, with theoretical continuum precision (although the implementation is limited to machine precision), the expected positions of the curves in the picture (see Fig. 1(b)). The construction of the potential distribution thus effects a transition from a local to a global representation of structure and shape.

Step 2.2. Curve Dynamics. The valleys of the potential distribution are located by dynamic unit-length curves, born at each tangent position in the tangent field. They evolve according to a variational scheme that depends on curve properties (tension and rigidity) as well as on the potential distribution. The evolution takes two forms: (i) a migration in position to achieve smooth coverings at subpixel accuracy; and (ii) a growth in length such that nearby covering elements overlap; see Fig. 2. At the end of the migration process, we shall have that each valley is completely covered by this collection of curves, and each curve will have deformed itself to fit the exact shape of the valley. We call this collection of curves the *global covering* of the curves in the image.

The dynamic curves of the second stage are a generalization of “snakes” [Kass, Witkin and Terzopoulos, 1988]. The key differences between our approach and theirs are twofold: (i) our scheme is parallel, so questions about total length, boundary points, etc., are handled implicitly; and (ii) our scheme is constructive, with the potential distribution computed directly from the tangent field, so questions about features, degree of image blurring, etc., are handled explicitly. We must neither have a user supply information interactively, as e.g. Kass et al. were required to do, nor, to avoid smoothing across discontinuities, do we have to distribute penalties globally, as e.g. Blake and Zisserman were required to do. Imposing stable intermediate structures—the tangent field and the potential distribution (see Fig. 1)—between the image and the global curves was the key, and the mathematical notion of coverings provided the unifying structure.

We now proceed to develop the notions of potential distribution and dynamic curves, before putting them together into global dynamic coverings in Sec. 5.

3. The potential distribution

The local descriptors in the tangent field are synthesized into a global description represented as a potential distribution. In this Section we first articulate this construction process, then define a mathematical notion of a valley in the potential distribution. Since valleys are intended to be the locus of points that represents (the trace of) a curve, finally we show that, for simple potential distributions, these valleys do indeed correspond to the global curves.

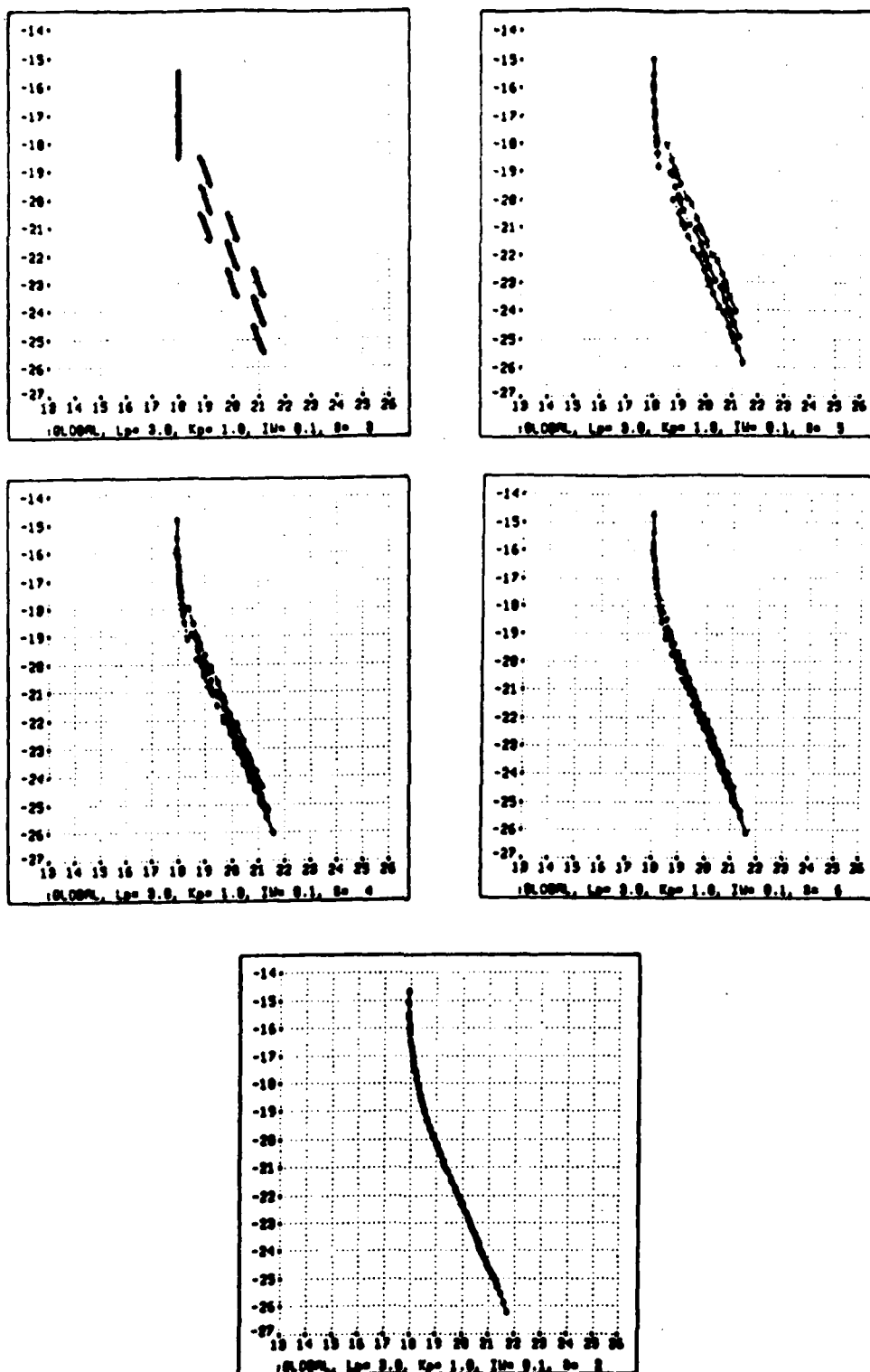


Figure 2 Illustration of how the collection of short dynamic curves evolve to cover the locus of points comprising the valleys in the potential distribution. In the usual order, the covering elements migrate from their initial positions (top left) to reach the bottom of the valley indicating the edge (bottom). Note the two aspects of the movement: a migration, and a length increase.

3.1 Constructing the Potential Distribution

The tangent field, consisting of discrete trace, discrete tangent, and discrete curvature, gives, at each trace point, a coarse estimate of the local behavior of curves. Since each tangent field entry corresponds not to a single curve in the world, but rather represents the equivalence class of possible curves that would project onto this tangent field entry, we represent this (local) equivalence class of possibilities by the (Gaussian-weighted) Wiener measure over them. The Wiener measure (a Gaussian distribution) arises because the class of continuous (but not necessarily differentiable) functions is equivalent to the sample functions of a Brownian motion, and the Gaussian weighting along it is an approximation to the diffusion of this measure in time [Doob, 1984]. The result, then, is a two-dimensional Gaussian distribution, and, since each entry in the tangent field is considered independently, the total potential distribution is the pointwise summation of them.

More formally, the set of generators of the potential field is the set of 2D-Gaussians $\{G_i : i = 1, \dots, n\}$. For each trace point i , G_i is obtained by the multiplication of two 1D-Gaussians, one in the direction of the quantized tangent at this point, and another one (with smaller deviation) in the perpendicular direction. Thus, at a trace point $i = (x_i, y_i)$, with tangent θ_i

$$\begin{aligned} G_i(x, y) &= - \left(K_E e^{-(f_i(x, y) - x_i)^2 / \sigma_E^2} \right) \left(K_B e^{-(g_i(x, y) - y_i)^2 / \sigma_B^2} \right) \\ &= -K_E K_B e^{-((f_i(x, y) - x_i)^2 / \sigma_E^2)} e^{-((g_i(x, y) - y_i)^2 / \sigma_B^2)} \end{aligned} \quad (1)$$

where the functions $f_i(x, y)$ and $g_i(x, y)$, which rotate the x and y axes to the axes indicated by the θ_i and $\theta_i + \frac{\pi}{2}$ directions respectively, are given by

$$\begin{aligned} f_i(x, y) &= x_i + (x - x_i) \cos \theta_i + (y - y_i) \sin \theta_i \\ g_i(x, y) &= y_i - (x - x_i) \sin \theta_i + (y - y_i) \cos \theta_i. \end{aligned}$$

The parameter σ_E , which is the standard deviation of the 1D-Gaussian in the θ_i -direction, controls how far the local information given by the tangent θ_i is extended from the trace point (x_i, y_i) . The parameter σ_B , which is the standard deviation of the 1D-Gaussian in the $(\theta_i + \frac{\pi}{2})$ -direction, controls the width over which the extended information is valid. For these reasons, we shall refer to the parameters σ_E and σ_B as the *extension parameter* and the *blurring parameter*, respectively, and K_E and K_B are (positive) weights associated with each of the 1D-Gaussians. Figure 3 presents a plot of the resulting 2D-Gaussian.

The (*tangent field*) *potential distribution*, U_{TF} , is then the pointwise summation of the G_i :

$$U_{TF}(x, y) = \sum_{i=1}^n G_i(x, y).$$

3. The potential distribution

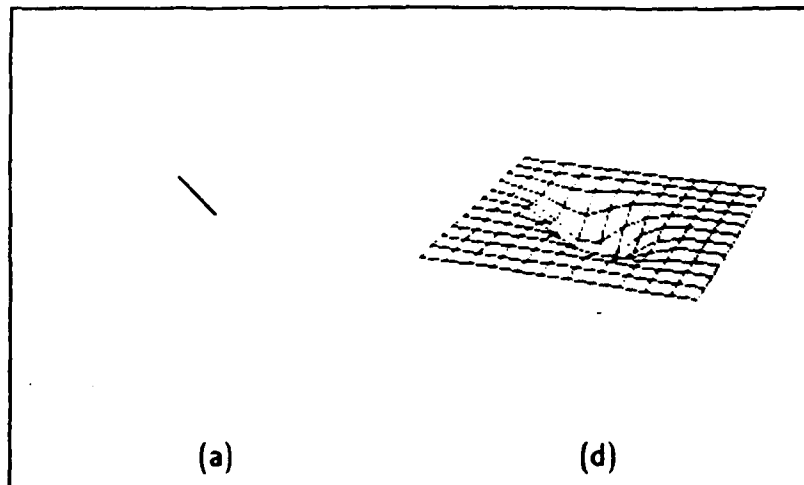


Figure 3 3D plot of $G_i(x, y)$. Note that the resulting Gaussian is elongated (in the θ_i -direction) since $\sigma_E > \sigma_B$.

It should be noted that each generator G_i , as defined, is a function of the tangent estimate alone, with the curvature estimate negligible. The curvature could be used if curved, 2D Gaussians G'_i were defined for each tangent field entry i . G'_i would then be a Gaussian along a curved axis of curvature κ_i in the θ_i direction, multiplied by another Gaussian perpendicular to the curved axis. However, the potential distribution U'_{TF} obtained by pointwise summation of these G'_i , which is

$$U'_{TF}(x, y) = \sum_{i=1}^n G'_i(x, y) ,$$

would be very similar to U_{TF} for the following reasons. First, the curvature is only retrieved in a very coarsely quantized form by the first stage of the curve detection process, so the generators G'_i would be only weakly tuned to curvature. Second, extending and summing the tangent information implicitly gives curvature, which is a relation between neighboring tangents. Thus, theoretically there appears to be no need to use curved Gaussians, and experimentally, potential distributions computed with straight Gaussians yielded smooth curved valleys. Figure 4 shows examples of potential distributions and of these valleys, where one can see that the result is a smooth "landscape", in which the jaggies due to sampling, noise and quantization have been removed ¹.

Recall from the introduction that the key feature of the potential distribution is the location of the valleys, since it is to the valleys that the covering elements migrate. Thus we can now state the claims behind this paper: (i) *the valleys in the potential distribution correspond to smooth curves, and (ii) the set of quantized tangents of these smooth curves is precisely the tangent field used to construct the distribution*. Moreover, over all the possible sets of curves which would project onto this tangent field, the valleys of the potential distribution give the most probable ones, in the sense of Wiener measure.

The rest of this Section is devoted to an analysis of the potential distribution characteristics, in order to prove the above claims. This analysis also results in a design criterion for the blurring parameter, which insures that separate curves remain separate.

The analysis is conducted for C^1 (i.e. continuous with continuous first derivatives) and non-intersecting curves, however some of the tangent fields used in the experiments contain intersecting curves and corners (in particular, see Figure 8). Other research is currently in progress to deal with intersecting curves and tangent discontinuities, by extending the potential distribution to a branched potential distribution; it will be reported in a separate paper ².

¹ Further computational experiments in which curved (G') generators were used led to final results virtually indistinguishable from those shown in this paper.

² Even if no specific treatment is performed at intersecting points and tangent discontinuities, the present

3. The potential distribution

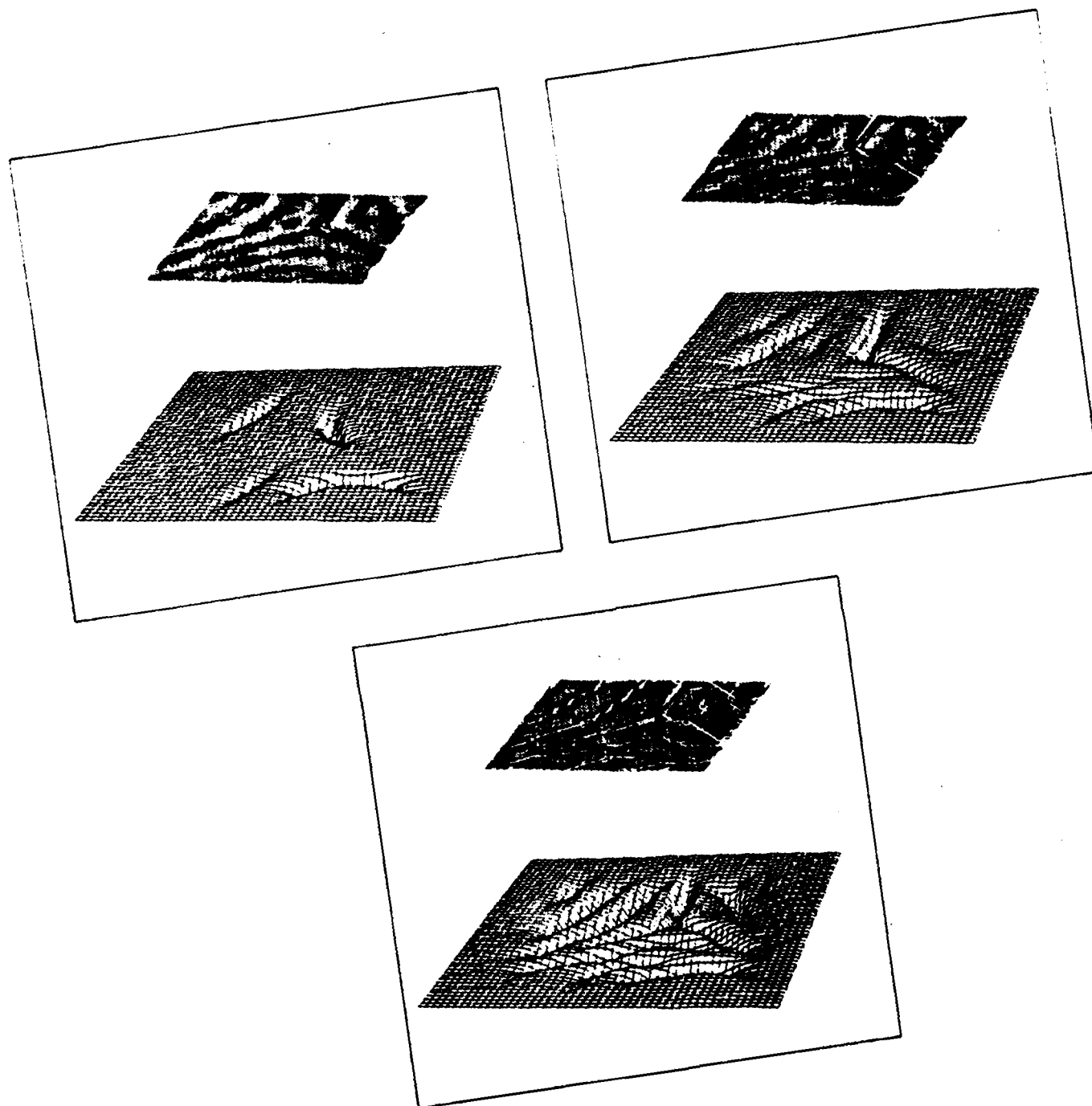


Figure 4 Examples of potential distributions. The image is the fingerprint image shown in Fig. 1. It is displayed in reverse contrast, with the tangent field shown in white. Three potential distributions are shown, computed from first a few, then more, and finally all the tangent field entries.

algorithm does not necessarily behave badly at these locations. For example, the T-shape in the middle

3.2 Valleys in the potential distribution

We now introduce the concept of a valley, the locus of points toward which the elements of our covering migrate. To begin, given a direction \mathbf{v} , making an angle ϕ with the positive x -axis, we define a ϕ -valley of the potential distribution $U_{TF}(x, y)$ to be a connected component of the set

$$\{(x, y) : \nabla_{(\phi+\frac{\pi}{2})} U_{TF}(x, y) = 0, \quad \nabla_{(\phi+\frac{\pi}{2})}^2 U_{TF}(x, y) > 0\},$$

where $\nabla_{\theta} f(x, y)$ and $\nabla_{\theta}^2 f(x, y)$ denote respectively the first and second directional derivative of f in the direction given by the angle θ . More generally, a valley of the potential distribution $U_{TF}(x, y)$ is a connected subset of the set

$$\{(x, y) : \nabla_{(\phi+\frac{\pi}{2})} U_{TF}(x, y) = 0, \quad \nabla_{(\phi+\frac{\pi}{2})}^2 U_{TF}(x, y) > 0\}$$

where $\phi = \phi(x, y)$, i.e. the angle ϕ can vary with the coordinates x and y . A ϕ -valley is then a "straight" valley in the direction ϕ .

The idea behind the concept of valleys is that of "directional minima". While local minima are points of a valley, points of a valley are not (necessarily) local minima of the potential distribution. The condition for a point (x, y) to be a local minimum of U_{TF} is that $\nabla_{\theta} U_{TF}(x, y) = 0$ and $\nabla_{\theta}^2 U_{TF}(x, y) > 0$ for all directions θ . The definition of a valley is thus weaker than the definition of a local minimum, and points in the valleys are minima with regard to a single direction, namely the direction given at each point by the angle $\phi(x, y)$.

3.3 Analysis of potential distributions

The analysis of the two simplest potential distributions, namely a single Gaussian, or perfectly aligned Gaussians, is simple and shows that the valleys are ϕ -valleys (where ϕ is the direction of the Gaussian) corresponding to the most probable smooth curve indicated by the tangent field. The formal treatment of those cases can be found in [David and Zucker, 1989].

In order to achieve this result for a more general potential distribution, we consider the case of two parallel Gaussians (this also leads to a criterion on the blurring parameter of the Gaussian weighting functions which insures curve separation).

Observe that, when the potential distribution is generated by several Gaussians that are in general position, we require the potential distribution to be such that: (1) Nearby tangent field entries which lie along the trace of a unique curve must form a unique valley, corresponding to the most probable position of the curve, and (2) Nearby tangent field

of the fingerprint image is correctly retrieved (see Figure 8). But further analysis would have to be done in order to insure the validity of the results in the general case.

entries belonging to two nearby curves must not be blurred together, and must give rise to two separate valleys.

These two requirements are sufficient to insure curve separation, since the curves are found by a migration of covering elements, born at tangent field entry positions, over the potential distribution (see Section 2.2). Then, in the case of tangent field entries belonging to different curves, the migration will stop as soon as the covering elements each drop into their surrounding valley. Hence, possible additional valleys between the two positions will never be found if two valleys exist sufficiently close to each one of the original positions.

We now concentrate on the second observation above, namely that nearly curves give rise to distinct valleys. This leads to our design criterion for σ_B . Again consider only two interacting Gaussians, but from a tangent field consisting of two parallel and vertical (i.e. $\theta_1 = \theta_2 = \frac{\pi}{2}$) tangents, with identical y -coordinates. Then, $U_{TF} = G_1(x, y) + G_2(x, y)$, and using $\theta_1 = \theta_2 = \frac{\pi}{2}$, $y_1 = y_2$ in Eq. 1, we get

$$\begin{aligned} G_1(x, y) &= -K_E K_B e^{-((y-y_1)^2/\sigma_E^2)} e^{-((x-x_1)^2/\sigma_B^2)} \\ G_2(x, y) &= -K_E K_B e^{-((y-y_1)^2/\sigma_E^2)} e^{-((x-x_2)^2/\sigma_B^2)} \end{aligned} \quad (2)$$

(see Fig. 5).

We now look for the $\frac{\pi}{2}$ -valleys of U_{TF} in this particular case. From now on, and without loss of generality, we assume $x_1 < x_2$. The critical points (for the $\frac{\pi}{2}$ -valleys) are those where $\nabla_x U_{TF} = 0$, or equivalently, those where $\frac{\partial}{\partial x} U_{TF}(x, y) = 0$. Computing the partial derivative, we obtain

$$\frac{2K_E K_B e^{-(y-y_1)^2/\sigma_E^2}}{\sigma_B^2} \left[(x-x_1) e^{-(x-x_1)^2/\sigma_B^2} + (x-x_2) e^{-(x-x_2)^2/\sigma_B^2} \right] = 0,$$

or, since first term of the product is always strictly positive,

$$(x-x_1) e^{-(x-x_1)^2/\sigma_B^2} + (x-x_2) e^{-(x-x_2)^2/\sigma_B^2} = 0. \quad (3)$$

Then, by positivity of the exponential function, any solution x of Equation 3 is in the open interval (x_1, x_2) . Also, Equation 3 depends only on $\Delta x = x_2 - x_1$ and σ_B . Without loss of generality, set $x_1 = 0$; then $x_2 = \Delta x$ and $F(x_1, x_2, \sigma_B)$ becomes

$$x e^{-x^2/\sigma_B^2} + (x - \Delta x) e^{-(x-\Delta x)^2/\sigma_B^2} \quad (4)$$

which depends only of Δx and σ_B , and which we denote $F(x)$ for short.

Observe that the points (x, y) belonging to the $\frac{\pi}{2}$ -valleys of U_{TF} are entirely characterized by the function $F(x)$ ¹. More precisely, it can be shown that (x_0, y_0) belongs to a $\frac{\pi}{2}$ -valley of U_{TF} if and only if $F(x_0) = 0$ and $\frac{\partial}{\partial x} F(x_0) > 0$.

¹ Note that this also means that the presence of the $\frac{\pi}{2}$ -valleys does not depend on the y coordinate.

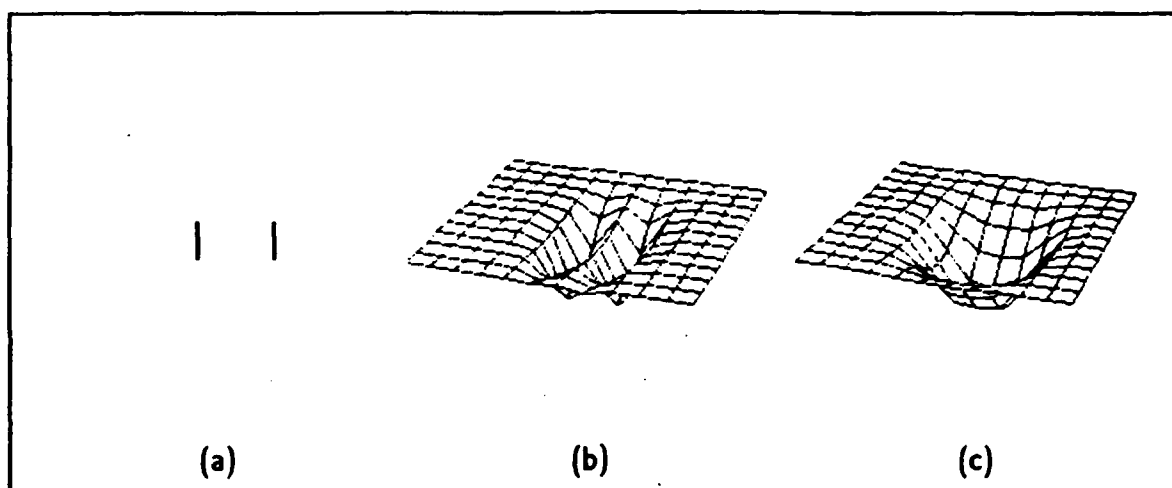


Figure 5 (a) The tangent field, consisting of two entries with $\theta_1 = \theta_2 = \frac{\pi}{2}$ and $v_1 = v_2$. (b) ... potential distribution obtained from (a) choosing $\sigma_E = 2.75$, $\sigma_B = \frac{\sigma_E}{4} = 0.6875$, $K_E = K_B = 1$. This Gaussian separates the tangents into two different valleys. (c) A potential distribution obtained from (a) choosing $\sigma_E = 2.75$, $\sigma_B = \frac{\sigma_E}{1.5} \approx 1.83$, $K_E = K_B = 1$. This Gaussian blurs the tangents into the same valley.

A condition to separate cases (b) and (c) of Figure 5 relates the distance Δx between the two Gaussian operators and the blurring parameter σ_B . More precisely, it can be deduced from last result [David and Zucker, 1989] that the vertical straight line $x = \frac{\Delta x}{2}$ is a $\frac{\pi}{2}$ -valley of U_{TF} if and only if $\Delta x^2 < 2\sigma_B^2$.

To summarize, the key structural properties of the potential distribution are:

- (i) When $\Delta x^2 < 2\sigma_B^2$, the valley at $x = \frac{\Delta x}{2}$ is unique.
- (ii) When $\Delta x^2 > 2\sigma_B^2$ so that $x = \frac{\Delta x}{2}$ becomes a maximum, two valleys are created, one at $x = x_1^* \approx x_1$, and one at $x = x_2^* \approx x_2$. Also, the valley at $x = \frac{\Delta x}{2}$ is lost. The slight variation in position from x_1 and x_2 arises because of a weak interaction between nearby Gaussians; see David and Zucker [1989].
- (iii) When Δx becomes large, the distances $(x_1^* - x_1)$ and $(x_2 - x_2^*)$ become small.
- (iv) When σ_B becomes small, the distances $(x_1^* - x_1)$ and $(x_2 - x_2^*)$ become small.

The preceding analysis thus indicates that the valleys in the potential distribution have precisely the properties required for synthesizing global curves from the tangent field. Before proceeding to the study of dynamic curves for locating these valleys, we extend one aspect of the preceding analysis into a constraint on the parameters controlling the Gaussian generators.

3.4 Design criterion: Blurring parameter

Continuing the particular case of preceeding Section, we know that two (vertical and parallel) tangent field entries give rise to a unique valley when $\Delta x^2 < 2\sigma_B^2$, and to two separate valleys when $\Delta x^2 > 2\sigma_B^2$. Recall from the introduction that image curves are represented in the tangent field as *connected* sets of tangent field entries. Thus, when two (parallel and vertical) tangent field entries each occupy overlapping (3x3) neighborhoods, they belong to the same curve, and when they are separated by at least one empty pixel, they belong to different curves¹. Then, any value of the blurring parameter such that

$$1^2 < 2\sigma_B^2 < 2^2 \quad \text{or} \quad \frac{1}{\sqrt{2}} < \sigma_B < \sqrt{2} \quad (5)$$

insures that the two tangent field entries give rise to separate valleys if and only if they belong to (locally) separate curves.

¹ Or, they belong to different parts of the same global curve, i.e. there is a chain somewhere between the two tangents. But *locally*, they must be viewed as separated.

More generally, it is clear that many other relations (in position and orientation), apart from the particular cases just studied, are possible, and lead to similar inequalities. For example, consider two parallel tangent field entries with $\theta = \frac{\pi}{4}$. These entries are then located at diagonal pixels and this case is just a rotation of the preceding one. But now, the distance between two neighboring tangent field entries is the distance across the diagonal, i.e. $\sqrt{2}$. Thus, any value of σ_B such that

$$\sqrt{2}^2 < 2\sigma_B^2 < (2\sqrt{2})^2 \quad \text{or} \quad 1 < \sigma_B < 2 \quad (6)$$

insures separation of (locally) disconnected sets of tangent field entries, and only of those sets. We submit that these two cases are typical of the general one, and summarize the analysis by saying that when

- (i) the analysis is restricted to non-intersecting curves;
- (ii) curves of the image are represented as connected sets of tangent field entries;
- (iii) the choice of σ_B is made according to Equations 5 and 6;

we obtain that

- (1) Each connected component of the tangent field will give rise to a unique valley, built from the interaction of all the tangent field entries of the component; thus, the valley indicates the position of the most natural curve which projects onto the connected component of tangent field entries.
- (2) Different connected components of the tangent field will be separated into distinct valleys. These valleys are not completely independent since there is a weak attraction between nearby valleys [David and Zucker, 1989]. But this attraction is non-negligible only for very close curves (a few pixels).

Experimentally, it will be shown in Section 6 that any reasonable value of σ_B relative to Equations 5 and 6 (say between 0.75 and 1.5) ensures that the two above claims are verified.

4. Valley detection

From the tangent field, we derived a potential distribution in which the valleys indicate the position of the curves in the image. We are now ready to identify these valleys to obtain a precise description of the curves. Since the valleys are, by definition, connected loci of points

that are not all minima of the potential distribution, standard minimization techniques are not adequate. Rather, we use short segments of curves to capture these loci, and we define a process by which the curves migrate into the valley. Together these segments will define the global covering. Figuratively, the process works as follows. Suppose that we are given a rough approximation of where a valley lies in the potential distribution. Now, if a curve is placed at this approximate position, within a potential well and parallel to the valley, and let slide as if it were governed by a kind of gravity, then it will slide toward the bottom of the well; i.e., the valley. Moreover, if the curve is not completely rigid, it will deform itself to fit the exact shape of the valley. Then, at the end of the process, when "gravitational" stability is attained, a description of the valley is given by the associated curve. This is precisely what we shall do.

Note that the above mechanism relies on the fact that we have a good initial approximation, in position and orientation, of each valley. If we place a piece of curve randomly on the potential distribution and let it evolve under "gravity", the final position of the curve could be meaningless in terms of valley detection. More likely than not, the curve would span several valleys. Our initial approximation is given by the tangent field, and the wells by the Gaussian generators. It is analyzed more carefully in Sec. 5. In this Section, we develop the variational principles that govern the dynamics of each deformable curve individually; the global cover will then be computed by running all of these dynamical processes in parallel. For now, we focus on one element of the cover, and suppose it exists at some (initial) position on the potential distribution. We then develop a mathematical model to govern the movement of the curve, basically following that adopted for *snakes* ([Kass, Witkin and Terzopoulos, 1987]; [Terzopoulos, 1987a and 1987b]), by applying classical mechanics to deformable curves.

A deformable curve v is a differentiable map $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. We write $v(s, t) = (x(s, t), y(s, t))$ with $s \in \Omega$ the space domain, and $t \in [t_0, t_1]$ the time domain. We denote $\frac{\partial}{\partial s}v(s, t)$ by $v_s(s, t)$, $\frac{\partial^2}{\partial s^2}v(s, t)$ by $v_{ss}(s, t)$, and $\frac{\partial}{\partial t}v(s, t)$ by $v_t(s, t)$.

We now define the *potential functional* $U^t(v)$ of a curve v , where the superscript indicates that the potential functional is computed at each time t . This corresponds to the potential energy of the curve. It is related to the potential distribution, but also includes internal constraints on the curve (how the curve is allowed to extend and fold). These internal constraints, corresponding to the physical concepts of tension and rigidity, must be added in order to specify completely the movement of the curve. The potential functional of a curve v at each time t is

$$U^t(v) = \frac{1}{2} \int_{\Omega} f(|v_s|^2) + g(|v_{ss}|^2) + U_{TF}(v) ds \quad (7)$$

where $U_{TF}(v)$ is the value of the potential distribution at the point $(x(s, t), y(s, t))$, $f(|v_s|^2)$ controls the tension of the curve, and $g(|v_{ss}|^2)$ controls the rigidity of the curve. Among all

the possible functions f and g , we simply choose the natural ones

$$\begin{aligned} f(|v_s(s, t)|^2) &= \omega_1(s)|v_s(s, t)|^2 \\ f(|v_{ss}(s, t)|^2) &= \omega_2(s)|v_{ss}(s, t)|^2 \end{aligned}$$

where ω_1 and ω_2 are constants over time. A method for choosing the constants is described in Sec. 5; we will also see that these simple forms are sufficient to control the curve movement for our purposes.

Now that the potential functional is completely defined, we can use the methods of classical mechanics (principle of least action, Lagrange-Euler equations of motion) and obtain a system of differential equations which describe the movement of the deformable curves. Then, with known initial positions, these equations can be solved numerically. All details can be found in [Terzopoulos, 1987a], [Terzopoulos, 1987b] and in the Appendix to this paper.

5. The global covering

We now proceed to put all of the previous pieces together. We infer the global curves from the tangent field by recovering a *covering* of these curves. Each element of the covering is a deformable curve that moves according to the model described in Sec. 4, within the wells of the potential distribution described in Sec. 3. Of course, we require good initial positions for the covering elements, so that, at “gravitational” stability, the curves will all have moved into valleys, and we require that each valley be completely covered by curves. In this Section, we precisely state what we mean by a covering of the global curves, and then develop a particular covering to meet the above two requirements.

5.1 Curves and coverings

As a mathematical object, a plane curve v is a mapping from an interval $I \subset \mathbb{R}$ to \mathbb{R}^2 , i.e.

$$v(s) = (x(s), y(s)) \quad \text{for } s \in I.$$

The set $v(I) \subset \mathbb{R}^2$ is called the *trace* of the curve v .

Since our overall goal is applications in vision, we seek to recover the trace of each curve as precisely as possible. Once this trace is available, it provides a basis for inferring all the other curve properties that do not depend on the exact form of the mapping. This leads us to consider equivalence classes of curves obtained by factoring the set of curves by the equivalence relation given by the equality of trace. More formally, we propose that two curves $v : I_v \rightarrow \mathbb{R}^2$ and $u : I_u \rightarrow \mathbb{R}^2$ are said to be *equivalent* if and only if their traces coincide i.e. $v(I_v) = u(I_u)$. But an image rarely consists of a unique curve, and we want to study general classes of images. We thus say that two sets of curves U and V , where U

consists of the curves $u : I_u \rightarrow \mathbb{R}^2$ and V consists of the curves $v : I_v \rightarrow \mathbb{R}^2$, are said to be *equivalent* if and only if

$$\bigcup_{u \in U} u(I_u) = \bigcup_{v \in V} v(I_v).$$

Given a set of curves V , we denote the equivalence class formed by all sets of curves equivalent to V by $E(V)$, and any member U of $E(V)$ is called a *covering* of the set V , or is said to cover V . The elements of U (which are curves $u : I_u \rightarrow \mathbb{R}^2$) are called *covering elements*.

5.2 Construction of the global covering

We now develop one covering which enjoys natural representational properties. Recall that reliable local estimates of the curve positions and orientations are given in the tangent field. Now, consider each tangent field entry as a unit-length straight curve, and let these n curves deform themselves according to the model already described. We then get the set C of these n deformed curves, where each curve $c \in C$ originated from a tangent field entry. In this Section, we will show that the set C is a covering (according to the above definition) of the global curves in an image. Or equivalently, since the global curves are described by the valleys of the potential field, we will show that each small curve c lies in a valley, and that each valley is completely covered by curves of the set C . More formally, the *global covering* of the curves in an image is the set C of the n curves

$$c_i(x, t_1) = (x_i(s, t_1), y_i(s, t_1)) \quad i = 1, \dots, n,$$

where at initial time t_0 , each $c_i(s, t_0)$ is a straight line identical to the i^{th} entry of the tangent field, and at final time t_1 , each $c_i(x, t_1)$ is a dynamic curve that has moved onto a valley of the potential distribution. Starting from now, the phrase "covering elements" will always refer to the elements of the global covering. We now complete their dynamics.

5.3 Covering element movement

Recall from the preceeding Section that for each covering element $c(s, t) = (x(s, t), y(s, t))$, with $s \in \Omega$, the potential functional is

$$\begin{aligned} U^t(c) &= \int_{\Omega} \omega_1 |c_s|^2 + \omega_2 |c_{ss}|^2 + U_{TF}(c) ds \\ &= \int_{\Omega} U_{TF}(c) ds + \int_{\Omega} \omega_1 |c_s|^2 + \omega_2 |c_{ss}|^2 ds \end{aligned} \quad (8)$$

where the first term controls the migration of the covering elements toward the valleys, and the second term controls how the covering elements are allowed to deform themselves during the migration. We call this latter term the *internal potential*, and set it so that each covering

element increases its length to a prescribed length. The movement of the covering elements is thus a two-fold process, consisting of a migration and of a length increase (see Fig. 2). The details are as follows.

The internal potential of a covering element $c(s, t) = (x(s, t), y(s, t))$, with $s \in \Omega$ is

$$\int_{\Omega} \omega_1 |c_s|^2 + \omega_2 |c_{ss}|^2 ds, \quad (9)$$

where ω_1 and ω_2 control tension and rigidity respectively. In fact, since the minimization of (9) participates in the covering elements' movement, we observe that the covering elements tend to decrease their length for positive values of ω_1 , increase their length for negative values of ω_1 , become straight for positive values of ω_2 , and fold and curve themselves for negative values of ω_2 .

Given a prescribed mandatory length L_p , we set ω_1 to be, at each time t , $\omega_1(c) = f(L^t(c) - L_p)$, where $L^t(c)$ is the length of the covering element $c(s, t)$ at time t , i.e.

$$L^t(c) = \int_{\Omega} |c_s(s, t)| ds = \int_{\Omega} \sqrt{x_s(s, t)^2 + y_s(s, t)^2} ds.$$

The function f is then simply a ramp function saturating when $|L^t(c) - L_p|$ becomes big. It is easy to see that this choice achieves the desired result: a covering element c will tend to extend when $L^t(c) < L_p$, and to shrink when $L^t(c) > L_p$. It should be noted that ω_1 is independent of the parameter s , which makes the tension equally distributed over the entire covering element length.

Although the presence of a rigidity term is essential to the process, its exact value is not a key point, and it suffices to set it empirically within an order of magnitude to get good results. It was experimentally verified that changing the rigidity parameter by a reasonable amount does not change the results significantly (see [David and Zucker, 1989]). But the rigidity parameter is essential to the algorithm, in order to prevent a covering element from folding on itself while finding its optimal position, since the potential distribution and the tension parameter do not prevent those movements. We simply set the rigidity parameter ω_2 to a constant value, the same for each covering element, at all times, and over the whole covering element length.

5.4 Properties of the global covering

Now that our model is completely defined, we can state and justify the claims made about the global covering. The global covering C , formed by the n curves c_i which originated from the n tangent field entries, and which evolve according to Equation 8, is such that

- (1) Each curve lies in a valley of the potential distribution.
- (2) Each valley of the potential distribution is completely covered by curves of the set C .

Claim (1) is met by construction, since the potential distribution was built from the tangent field, i.e. from the curves' initial positions. Then, each curve is, at initial time, within the well surrounding its associated valley, and the dynamics drive the curve onto the valley.

The truth of claim (2) is based upon three facts: global curves project into the tangent field as connected sets of tangent field entries, the length of each covering element increases during the dynamics, and there is no movement of a covering element outside the well surrounding its valley. By the last statement, we assume that all valleys have a similar depth along their length, or equivalently, that the density of the tangent field entries is about the same along the entire connected component. This is guaranteed by our restriction to smooth non-intersecting curves.

Note that the diagonal length across a pixel (which is the maximal distance between two neighboring elements of a connected set) is $\sqrt{2}$ units; therefore, to guarantee that neighboring covering elements overlap to give the global curves, we just have to let each covering element extend in length to some value greater than $\sqrt{2}$, since each covering element stays around its initial position. Typically, a prescribed length of 3 is used, in order to guarantee the overlap.

Since the valleys of the potential field describe the global curves of the image, we also get that the global covering is a covering of the curves in an image, i.e.

$$\bigcup_{c \in C} c(\Omega, t_1) = \bigcup_{v \in V} v(I_v),$$

where V is the set of the global curves of the image.

6. Results

We now illustrate the computation of the global covering, beginning with artificial images to check the validity of the algorithm against known examples, and finally with natural images to demonstrate its robustness.

The first test shows that simple analytic curves are retrieved quite exactly from the tangent field information, i.e. that the "natural" curves indicated by the potential distribution correspond to the expected ones. The image consists of two nearby circles, with different radii, projected over a discrete sampling grid (Fig. 6(a)).

We simulate the first stage of the process (see Sec.2.1), which extracts the trace and tangent of these curves in the following manner. Each square of the sampling grid traversed by one of the circles belongs to the trace of this circle, and the tangent at each trace point is computed assuming that the curve crosses the center of the square; afterwards, all the computed tangents are quantized into eight classes (Fig. 6(b)). We then run our algorithm on the simulated tangent field. This tangent field is the only information that the second

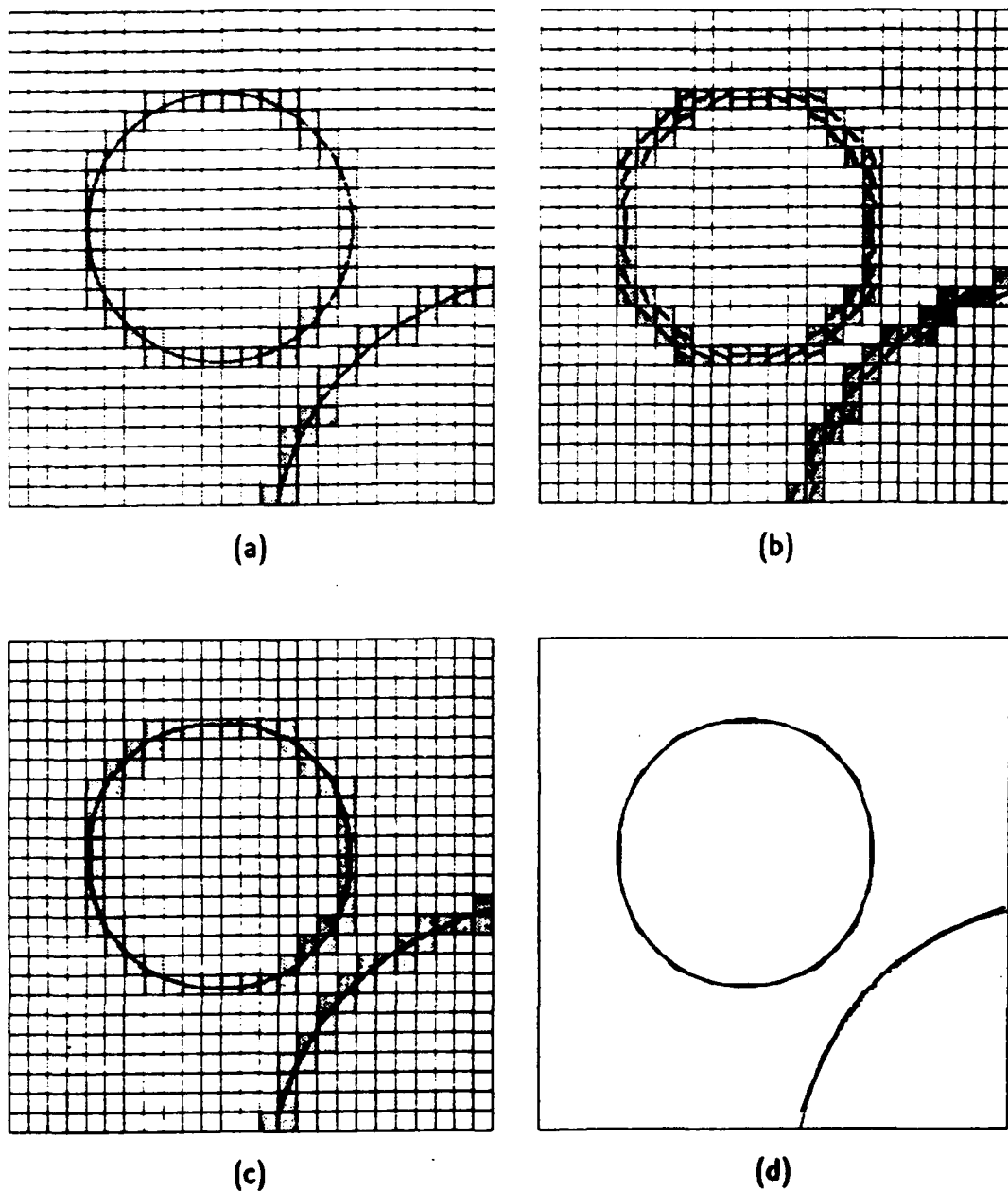


Figure 6 (a) Two nearby circles of different radii lie over a discrete sampling grid (25 x 25). Their projection onto the grid is given by those pixels traversed by one of the two circles. (b) The tangent field for the two circles, consisting of the discretized tangent at each trace point. Note that the tangents' positions also correspond to the covering elements initial positions. (c) From their initial positions shown in (b), the covering elements have migrated into the valleys of the potential distribution, where they attain stability. The parameters of this experiment are: $\sigma_E = 2.75$, $\sigma_B = \sigma_E/3 = 0.58$, $L_P = 3.0$; ω_1 is determined by $R = 1.0$, $K = 1.0$, and $\omega_2 = 1.0$; $\mu = 2.0$, $\gamma = 2.0$. (d) The final covering elements' positions, superimposed over the initial curves.

stage of the process uses to retrieve the global curves, the initial circles being lost by the projection onto the sampling grid. The result is shown in Fig. 6(c) and (d). The curves retrieved from the potential distribution correspond very accurately to the initial circles. Also, the value of σ_B , chosen according to the design criterion, allows a clear separation between the two nearby circles, separated by only two pixels on the sampling grid.

Finally, we run our algorithm on three natural images, a fingerprint image, a radiograph of blood vessels in the brain, and a satellite image of logging roads. Since we are interested in precise extraction of the curves, without blurring of very close curves, we choose to experiment with small images so that details could be examined up close. Figure 7 presents the original images, and the resulting global curves. An exploded display of the first two tests is then presented in the last two figures.

7. Summary and Conclusions

In this paper we developed an algorithm for finding a representation of the global curves in an image. The novel feature of this representation is that it is formulated in mathematical terms as a covering, so that, rather than computing global curves directly, short segments of curve are computed independently from one another. The short segments thus become the elements of the cover, and global curves are given as their union. This enables us to define the algorithm in a parallel fashion, and to avoid the *a priori* specification of global parameters (e.g., the length of the curve).

Since the global representation is built up from local ones, the algorithm for computing global curves via coverings requires reliable local information about the curve. Within our context, which is a large project on curve detection incorporating both mathematical and biological constraints, this local information is provided by a tangent field, or a list of (quantized trace) points through which the curve passes, together with coarse estimates of its tangent and curvature at those points. The key point to emphasize here is that these local properties need not be specified to high precision; the interpolation properties implicit within the covering algorithm can deal successfully with coarse initial information. Thus the algorithm fits into a very natural information processing hierarchy, with both degree-of-precision and local-to-global axes spanned.

The local (tangent field) information is used in two different ways. First, a global description, in the form of a potential distribution, is synthesized from it. By definition, valleys in this potential distribution correspond precisely to curve locations. These valleys are detected by a dynamical process in which the covering elements are driven by the potential to both migrate in position and elongate in length. The second use of the local (tangent field) information is in specifying the initial positions for each covering element. Convergence is thereby guaranteed by construction. Equivalence between the curves in the image and the covering elements is demonstrated both theoretically and by example.

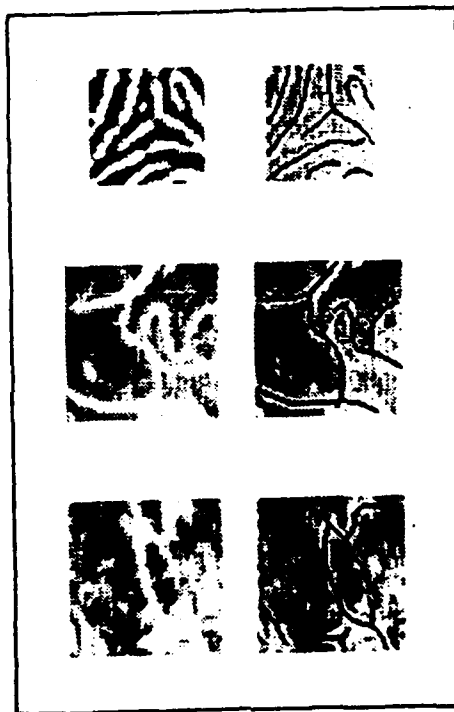


Figure 7 Global curves of natural images. Top: A fingerprint image. Middle: A radiograph of blood vessels in the brain. Bottom: A satellite image of logging roads.



Figure 8 Illustration of the first experiment on a fingerprint image. (a) The initial covering elements, at tangent field entries' positions. (b) Final positions of the covering elements; clearly the covering elements have migrated into the valleys, and have overlapped to form a covering. Also note that the bifurcation in the center of the image have been recovered correctly, without special treatment.

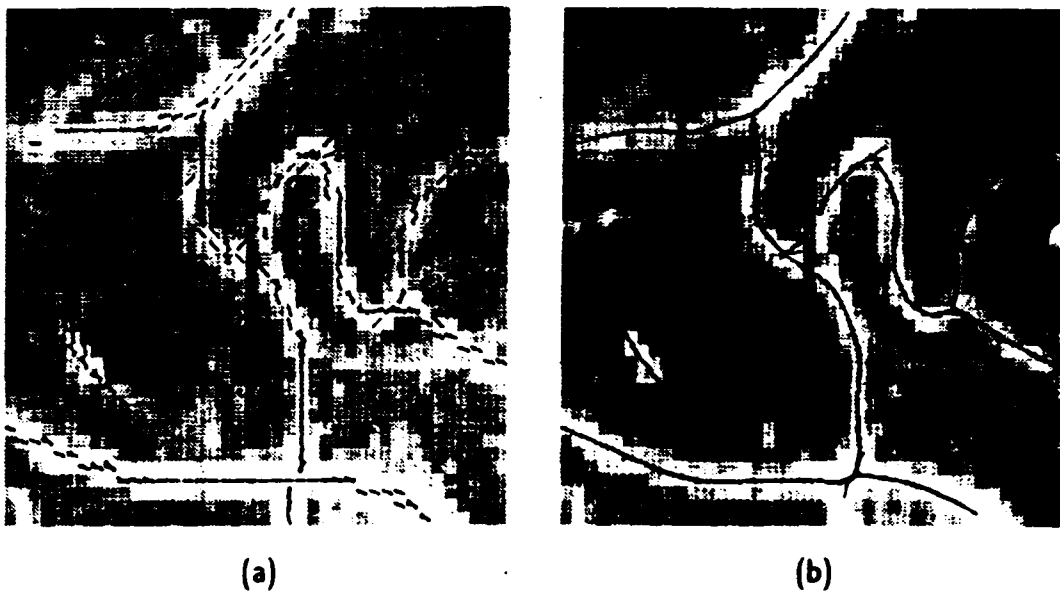


Figure 9 The second test is performed on a biomedical image, a cerebral angiogram. (a) The initial positions of the covering elements, corresponding to the tangent field entries. This image is very noisy, and low contrast. The tangent field was computed under the assumption that all vessels were of comparable contrast. (b) The final positions of the covering elements. The region where nearby covering elements do not overlap completely corresponds to a region where the tangent field was not completely connected.

Since the global curves are represented by a covering that evolves according to a dynamic process, we refer to them as *dynamic global coverings*. Although we concentrated on smooth, non-intersecting curves in this paper, additional research is in progress to guarantee the validity of the results in the general case. It works as follows. Since the tangent field indicates curve intersections and discontinuities by multiple tangent field entries at the same grid coordinate, the potential distribution is split into several "layers", one per tangent field entry, in the neighborhood of each of those coordinates. Take, for example, the case of two curves intersecting at a single point. The tangent field entries of those curves will give rise to two independent layers in the potential distribution at this point, and the process described in this paper can then be applied at each layer. (The advantage of separating the tangent field entries is, of course, to eliminate inappropriate interactions between the separate curve segments.)

The remaining step from a covering to a global representation is a specification of which covering elements belong to the same global curve. (It is usually the case, for example, that there are multiple curves within a single image.) We accomplish this by a straightforward exploitation of topological connectivity: graphically, the initial covering elements are all born with a different "color"; as two covering elements overlap dynamically, they become the same "color". In the end, "color" propagates along connected components, and the elements belonging to the covering of the same curve are all the same color, while the covers corresponding to different curves are different colors.

In conclusion, we believe that dynamic coverings have application to a variety of problems that must face the transition from local to global representations in a parallel, efficient way. For example, an extension to surface coverings could work as follows. We know that it is possible to extract, from 3D data, an intermediate structure that is the 3D equivalent of the tangent field (see Sander and Zucker [1989]). This structure mainly consists of an estimate of the tangent plane at each (3D) trace point. Starting from these planes, a generalization of the model presented in this paper would evolve those local estimates to fit together into a (smooth) covering of the surface. The covering elements would thus be overlapping surface patches.

8. References

- Ballard, B., and Brown, C., *Computer Vision*, Prentice Hall, Englewood Cliffs, 1982.
- Benson, A., and Evans, D.J., A normalized algorithm for the solution of positive definite symmetric quintadiagonal systems of linear equations, *ACM Trans. Mathematical Software*, **3**, 1977, 96-103.
- Blake, A., and Zisserman, A., *Visual Reconstruction*, The MIT Press, Cambridge, Massachusetts, 1987.

- David, C., and Zucker, S.W., Potential, Valleys, and Dynamic Global Coverings, Technical Report TR-CIM 89-1, McGill University, Montréal, 1989.
- Dobbins, A., Zucker, S.W., and Cynader, M.S., Endstopping in the visual cortex as a substrate for calculating curvature, *Nature*, **329**, 438-441, 1987.
- Doob, J.L., *Classical Potential Theory and Its Probabilistic Counterpart* Springer Verlag, New York, 1984.
- Draper, B., Collins, R., Brolio, J., Hanson, A., and Riseman, E., The schema system, *Int. J. Computer Vision*, 1989, to be published.
- Hummel, R., and Zucker, S.W., On the foundations of relaxation labelling processes, *IEEE Trans. PAMI*, **5**, 1983, 267-287.
- Iverson, L., and Zucker, S.W., From orientation selection to optical flow: a computational perspective, *Proc. IEEE Workshop in Computer Vision*, Miami, 1987. 1987, 184-189.
- Iverson, L.A., The description of images curves: Discrete forms of continuity, M. Eng. Thesis, McGill University, Montréal, 1988.
- Kaas, W., and Witkin, A., Analyzing oriented patterns, *Computer Vision, Graphics and Image Processing*, **37**, 1987, 362-385.
- Kaas, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, *Proc. ICCV*, **1**, 259-268.
- Kaplan, W., *Advanced Calculus*, Addison-Wesley, 1952.
- Levine, M., *Vision in Man and Machine*, McGraw Hill, New York, 1985.
- Lowe, D.G., Organization of smooth image curves at multiple scales, *Proc. ICCV*, **2**, 1988, 558-567.
- Montanari, U., On the optimal detection of curves in noisy pictures, *Comm. ACM*, **14**, 1971, 335-345.
- Parent, P., and Zucker, S.W., Trace inference, curvature consistency, and curve detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1989, **11**, 823 - 839.
- Sander, P., and Zucker, S.W., Inferring differential structure from 3-D images: Smooth cross sections of fibre bundles, *IEEE Trans. PAMI*, to be published.
- Terzopoulos, D., Image analysis using multigrid relaxation methods, *IEEE Trans. PAMI*,

1986. **PAMI-8**, 129 - 139.
- Terzopoulos, D., On matching deformable models to images, *Topical Meeting on Machine Vision*, Technical Digest Series, **12**, 1987, 160-163.
- Terzopoulos, D., On matching deformable models to images: Direct and iterative solutions. *Topical Meeting on Machine Vision*, Technical Digest Series, **12**, 1987, 164-167.
- Tsotsos, J., Image understanding, in *Encyclopedia of Artificial Intelligence*, (S. Shapiro Ed.). John Wiley, New York, 1987, 389-409.
- Zucker, S.W., Early orientation selection: Tangent fields and the dimensionality of their support, *Computer vision, Graphics and Image Processing*, **32**, 1985, 74-103.
- Zucker, S.W., The emerging paradigm of computational vision, in *Annual Reviews of Computer Science*, vol. 2, 1987, 69 - 89.
- Zucker, S.W., Dobbins, A., and Iverson, L., Two stages of curve detection suggest two styles of visual computation, *Neural Computation*, 1989, 1, 68 - 81.
- Zucker, S.W., David, C., Dobbins, A., Iverson, L., The organization of curve detection: Coarse tangent fields and fine spline coverings, *Proc. 2nd International Conference on Computer Vision*, Tarpon Springs, 1988.

APPENDIX

A.1 Dynamics of the deformable curve

Given the potential functional $U^t(v)$, we can define the kinetic energy of the curve. Following classical mechanics, the kinetic energy functional at each time t is

$$T^t(v) = \frac{1}{2} \int_{\Omega} \mu |v_t|^2 ds,$$

where μ is the (constant) mass density.

According to the *principle of least action*, the motion of the deformable curve during the time interval $[t_0, t_1]$ is described by the functions $x(s, t)$, $y(s, t)$ for which the integral

$$\int_{t_0}^{t_1} L^t(v) dt$$

of the Lagrangian $L^t(v) = T^t(v) - U^t(v)$ is a minimum. It is also known that any extrema of $\int_{t_0}^{t_1} L^t(v) dt$ must satisfy the Euler-Lagrange equations. Setting $f(v) = \frac{1}{2}(\mu|v_t|^2 - \omega_1|v_s|^2 - \omega_2|v_{ss}|^2 - U_{TF}(v))$, the Euler-Lagrange equations are

$$\begin{aligned} f_x - \frac{\partial}{\partial s}(f_{xs}) - \frac{\partial}{\partial t}(f_{xt}) + \frac{\partial^2}{\partial s^2}(f_{xss}) + \frac{\partial^2}{\partial t^2}(f_{xtt}) + \frac{\partial^2}{\partial s \partial t}(f_{xst}) &= 0 \\ f_y - \frac{\partial}{\partial s}(f_{ys}) - \frac{\partial}{\partial t}(f_{yt}) + \frac{\partial^2}{\partial s^2}(f_{yss}) + \frac{\partial^2}{\partial t^2}(f_{ytt}) + \frac{\partial^2}{\partial s \partial t}(f_{yst}) &= 0 \end{aligned} \quad (.1)$$

where, for example, f_{xst} denotes $\frac{\partial}{\partial x_{st}} f$, the derivative with respect to the variable x_{st} .

Equations .1 hold for a conservative system. In order to dissipate the kinetic energy produced during the motion, an energy dissipation functional can be introduced, such that the deformable curves reach stable equilibrium positions. Using the Raleigh dissipation functional $D^t(v) = \int_{\Omega} \gamma |v_t|^2 ds$, where γ is the (constant) damping density, (.1) becomes

$$\begin{aligned} f_x - \frac{\partial}{\partial s}(f_{xs}) - \frac{\partial}{\partial t}(f_{xt}) + \frac{\partial^2}{\partial s^2}(f_{xss}) + \frac{\partial^2}{\partial t^2}(f_{xtt}) + \frac{\partial^2}{\partial s \partial t}(f_{xst}) + \frac{\partial}{\partial x_t}(\gamma |v_t|^2) &= 0 \\ f_y - \frac{\partial}{\partial s}(f_{ys}) - \frac{\partial}{\partial t}(f_{yt}) + \frac{\partial^2}{\partial s^2}(f_{yss}) + \frac{\partial^2}{\partial t^2}(f_{ytt}) + \frac{\partial^2}{\partial s \partial t}(f_{yst}) + \frac{\partial}{\partial y_t}(\gamma |v_t|^2) &= 0. \end{aligned}$$

Letting $\Gamma_x(v)$ denote $\frac{\partial}{\partial x} U_{TF}(v)$, we calculate

$$\begin{aligned} f_x &= -\frac{1}{2}\Gamma_x(v) \\ \frac{\partial}{\partial s}(f_{xs}) &= -\frac{\partial}{\partial s}(\omega_1 x_s) \\ \frac{\partial^2}{\partial s^2}(f_{xss}) &= -\frac{\partial^2}{\partial s^2}(\omega_1 x_{ss}) \\ \frac{\partial}{\partial t}(f_{xt}) &= \mu x_{tt} \\ \frac{\partial^2}{\partial t^2}(f_{xtt}) &= \frac{\partial^2}{\partial s \partial t}(f_{xst}) = 0. \end{aligned}$$

Using identical results for f_y , $\frac{\partial}{\partial s}(f_{ys})$, $\frac{\partial^2}{\partial s^2}(f_{yss})$, $\frac{\partial}{\partial t}(f_{yt})$, $\frac{\partial^2}{\partial t^2}(f_{ytt})$ and $\frac{\partial^2}{\partial s \partial t}(f_{yst})$, the Lagrange equations of motion of the deformable curve are

$$\begin{aligned} \mu x_{tt} + \gamma x_t - \frac{\partial}{\partial s}(\omega_1 x_s) + \frac{\partial^2}{\partial s^2}(\omega_2 x_{ss}) &= -\frac{1}{2}\Gamma_x(v) \\ \mu y_{tt} + \gamma y_t - \frac{\partial}{\partial s}(\omega_1 y_s) + \frac{\partial^2}{\partial s^2}(\omega_2 y_{ss}) &= -\frac{1}{2}\Gamma_y(v). \end{aligned} \quad (.2)$$

A.2 Numerical solution

The solution of System .2 gives the positions, in space and in time, of the curve r during the motion. We now address the numerical approximation to System .2, following Terzopoulos [1988b].

The space domain Ω is tessellated into $N+1$ nodes $\{0, h, 2h, \dots, Nh\}$, $h = \frac{1}{N}$. Then, the solutions to the system are the vectors $\mathbf{x}^t = (x(ih, t))_{i=0}^N = (x_i^t)_{i=0}^N$, $\mathbf{y}^t = (y(ih, t))_{i=0}^N = (y_i^t)_{i=0}^N$ at each time t . It is possible to express (.2) as the linear system

$$A\mathbf{x}^t = \mathbf{g}_x^t, \quad A\mathbf{y}^t = \mathbf{g}_y^t.$$

Assuming temporarily that the curve is closed, two initial configurations (initial conditions) are needed. Solving for \mathbf{x}^t , given a time step Δt , x_t and x_{tt} are approximated by

$$x_t = \frac{x^t - x^{t-2\Delta t}}{2\Delta t}$$

$$x_{tt} = \frac{x^{t-2\Delta t} - 2x^{t-\Delta t} + x^t}{\Delta t^2},$$

where backward differences are used, since we can only rely on curve positions at previous times. This results in the system of $N+1$ equations

$$\begin{aligned} & -\frac{\partial}{\partial s}(\omega_1 x_s) + \frac{\partial^2}{\partial s^2}(\omega_2 x_{ss}) + \left(\frac{\mu}{\Delta t^2} + \frac{\gamma}{2\Delta t}\right)x_i^t \\ & = -\frac{1}{2}\Gamma_x(x_i^{t-\Delta t}, y_i^{t-\Delta t}) + 2\frac{\mu}{\Delta t^2}x_i^{t-\Delta t} - \left(\frac{\mu}{\Delta t^2} - \frac{\gamma}{2\Delta t}\right)x_i^{t-2\Delta t}, \end{aligned} \quad (.3)$$

$i = 0, \dots, N$

where $\frac{\partial}{\partial x}U_{TF}(v)$ is evaluated at time $t - \Delta t$.

The right hand sides of Equations .3 only depend on prior configurations, and can be evaluated at each time step t , as long as two initial configurations ($\mathbf{x}^{t_0-\Delta t}$ and $\mathbf{x}^{t_0-2\Delta t}$) are given at the initial time t_0 . Setting

$$\mathbf{g}_x^t = \left(-\frac{1}{2}\Gamma_x(x_i^{t-\Delta t}, y_i^{t-\Delta t}) + 2\frac{\mu}{\Delta t^2}x_i^{t-\Delta t} - \left(\frac{\mu}{\Delta t^2} - \frac{\gamma}{2\Delta t}\right)x_i^{t-2\Delta t} \right)_{i=0}^N$$

and writing

$$A = \left(\frac{\mu}{\Delta t^2} + \frac{\gamma}{2\Delta t}\right)I + K = \alpha I + K,$$

(.3) reduces to

$$(\alpha I + K)\mathbf{x}^t = \mathbf{g}_x^t$$

where the only unknown is the *stiffness matrix* K determined by the relations between nodes at each time t . Again using finite differences to approximate the spatial derivatives $\frac{\partial}{\partial s}(\omega_1 x_s)$ and $\frac{\partial^2}{\partial s^2}(\omega_2 x_{ss})$ (see Section .4), we get

$$\begin{aligned} -\frac{\partial}{\partial s}(\omega_1 x_s) + \frac{\partial^2}{\partial s^2}(\omega_2 x_{ss}) \approx \\ x_{i-2h} \left[\frac{\omega_2(i-h)}{h^4} \right] \\ + x_{i-h} \left[\frac{-2\omega_2(i)}{h^4} - \frac{2\omega_2(i-h)}{h^4} - \frac{\omega_1(i-h)}{h^2} \right] \\ + x_i \left[\frac{\omega_2(i+h)}{h^4} + \frac{4\omega_2(i)}{h^4} + \frac{\omega_2(i-h)}{h^4} + \frac{\omega_1(i)}{2h^2} - \frac{\omega_1(i-h)}{h^2} \right] \\ + x_{i+h} \left[\frac{-2\omega_2(i+h)}{h^4} - \frac{2\omega_2(i)}{h^4} - \frac{\omega_1(i)}{h^2} \right] \\ + x_{i+2h} \left[\frac{\omega_2(i+h)}{h^4} \right] \end{aligned}$$

in the case of closed curves (i.e. $x_{-1}^t = x_N^t$, $x_{-2}^t = x_{N-1}^t$, $x_{N+1}^t = x_0^t$, $x_{N+2}^t = x_1^t$).

Finally, setting

$$\begin{aligned} a_i &= \frac{\omega_2(i+h)}{h^4} \\ b_i &= -\frac{2\omega_2(i)}{h^4} - \frac{2\omega_2(i+h)}{h^4} - \frac{\omega_1(i)}{h^2} \\ c_i &= \frac{\omega_2(i-h)}{h^4} + \frac{4\omega_2(i)}{h^4} + \frac{\omega_2(i+h)}{h^4} + \frac{\omega_1(i-h)}{h^2} + \frac{\omega_1(i)}{h^2} \end{aligned}$$

for $i = 1, \dots, N$, the stiffness matrix K is the pentadiagonal symmetric matrix

$$\begin{pmatrix} c_0 & b_0 & a_0 & & & & a_{N-1} & b_N \\ b_0 & c_1 & b_1 & a_1 & & & & a_N \\ a_0 & b_1 & c_2 & b_2 & a_2 & & & \\ & a_1 & b_2 & c_3 & b_3 & a_3 & & \\ \dots & & & & & & & \\ & & & a_{N-4} & b_{N-3} & c_{N-2} & b_{N-2} & a_{N-2} \\ a_{N-1} & & & a_{N-3} & b_{N-2} & c_{N-1} & d_{N-1} & \\ b_N & a_N & & & a_{N-2} & b_{N-1} & c_N & \end{pmatrix}.$$

Then, $A = \alpha I + K$ is also pentadiagonal symmetric and can be solved very efficiently in $O(N)$ time and space [Benson and Evans, 1977], by factoring A into triangular and diagonal matrices.

A.3 Inserting position discontinuities

Up to now, the deformable curve was assumed to be closed. We now relax this assumption by inserting a discontinuity between node $N+1$ and node 0.

To insert a position discontinuity at $s = (\frac{i-1}{2})h$ between nodes $i-1$ and i , the stiffness matrix K has to be changed to eliminate relations between nodes $i-1$ and i . The equations which relate nodes $i-1$ and i are given by the matrix rows $i-2$, $i-1$, i and $i+1$, and are the following

$$\begin{aligned}
 a_{i-4}x_{i-4} + b_{i-3}x_{i-3} + c_{i-2}x_{i-2} + b_{i-2}x_{i-1} + a_{i-2}x_i &= g_{i-2} \\
 a_{i-3}x_{i-3} + b_{i-2}x_{i-2} + c_{i-1}x_{i-1} + b_{i-1}x_i + a_{i-1}x_{i+1} &= g_{i-1} \\
 a_{i-2}x_{i-2} + b_{i-1}x_{i-1} + c_i x_i + b_i x_{i+1} + a_i x_{i+2} &= g_i \\
 a_{i-1}x_{i-1} + b_i x_i + c_{i+1}x_{i+1} + b_{i+1}x_{i+2} + a_{i+1}x_{i+3} &= g_{i+1}.
 \end{aligned} \tag{.4}$$

From (.4), we immediately get $b_{i-1} = a_{i-2} = a_{i-1} = 0$, which implies $\omega_2(i-1) = \omega_2(i) = \omega_1(i-1) = 0$. This dictates the 9 entries to change in the matrix, which are

1. $a_{i-2} = 0$
2. $a_{i-1} = 0$
3. $b_{i-2} = -\frac{2\omega_2(i-2h)}{h^4} - \frac{\omega_1(i-2h)}{h^2}$
4. $b_{i-1} = 0$
5. $b_i = -\frac{2\omega_2(i+h)}{h^4} - \frac{\omega_1(i)}{h^2}$
6. $c_{i-2} = \frac{\omega_2(i-3h)}{h^4} + \frac{\omega_2(i-2h)}{h^4} + \frac{\omega_1(i-3h)}{h^2} + \frac{\omega_1(i-2h)}{h^2}$
7. $c_{i-1} = \frac{\omega_2(i-2h)}{h^4} + \frac{\omega_1(i-2h)}{h^2}$
8. $c_i = \frac{\omega_2(i+h)}{h^4} + \frac{\omega_1(i)}{h^2}$
9. $c_{i+1} = \frac{4\omega_2(i+h)}{h^4} + \frac{\omega_2(i+2h)}{h^4} + \frac{\omega_1(i)}{h^2} + \frac{\omega_1(i+h)}{h^2}$.

In our implementation, this is used only once, to relax the condition that the curve must be a closed one.

A.4 Development of the spatial derivatives

We indicate here how the numerical approximation to spatial derivatives are obtained.

Centered, backward and forward differences are mixed in order to obtain a symmetric matrix.

$$\begin{aligned}
\frac{\partial}{\partial s}(\omega_1 x_s) &\approx \frac{\omega_1(i)x_s(i) - \omega_1(i-h)x_s(i-h)}{h} \\
&\approx \frac{\omega_1(i)}{h} \left[\frac{x_{i+h} - x_i}{h} \right] - \frac{\omega_1(i-h)}{h} \left[\frac{x_i - x_{i-h}}{h} \right] \\
&= x_{i-h} \left[\frac{\omega_1(i-h)}{h^2} \right] + x_i \left[\frac{-\omega_1(i-h) - \omega_1(i)}{h^2} \right] + x_{i+h} \left[\frac{\omega_1(i)}{h^2} \right] \\
\frac{\partial^2}{\partial s^2}(\omega_2 x_{ss}) &\approx \frac{\omega_2(i+h)x_{ss}(i+h) - 2\omega_2(i)x_{ss}(i) + \omega_2(i-h)x_{ss}(i-h)}{h^2} \\
&\approx \frac{\omega_2(i+h)}{h^2} \left[\frac{x_{i+2h} - 2x_{i+h} + x_i}{h^2} \right] \\
&\quad - \frac{2\omega_2(i)}{h^2} \left[\frac{x_{i+h} - 2x_i + x_{i-h}}{h^2} \right] \\
&\quad + \frac{\omega_2(i-h)}{h^2} \left[\frac{x_i - 2x_{i-h} + x_{i-2h}}{h^2} \right] \\
&= x_{i-2h} \left[\frac{\omega_2(i-h)}{h^4} \right] + x_{i-h} \left[\frac{-2\omega_2(i) - 2\omega_2(i-h)}{h^4} \right] \\
&\quad + x_i \left[\frac{\omega_2(i+h) + 4\omega_2(i) + \omega_2(i-h)}{h^4} \right] \\
&\quad + x_{i+h} \left[\frac{-2\omega_2(i+h) - 2\omega_2(i)}{h^4} \right] \\
&\quad + x_{i+2h} \left[\frac{\omega_2(i+h)}{h^4} \right]
\end{aligned}$$

Thus, $-\frac{\partial}{\partial s}(\omega_1 x_s) + \frac{\partial^2}{\partial s^2}(\omega_2 x_{ss})$ is numerically approximated by

$$\begin{aligned}
&x_{i-2h} \left[\frac{\omega_2(i-h)}{h^4} \right] \\
&+ x_{i-h} \left[\frac{-2\omega_2(i)}{h^4} - \frac{2\omega_2(i-h)}{h^4} - \frac{\omega_1(i-h)}{h^2} \right] \\
&+ x_i \left[\frac{\omega_2(i+h)}{h^4} + \frac{4\omega_2(i)}{h^4} + \frac{\omega_2(i-h)}{h^4} + \frac{\omega_1(i)}{2h^2} + \frac{\omega_1(i-h)}{h^2} \right] \\
&+ x_{i+h} \left[\frac{-2\omega_2(i+h)}{h^4} - \frac{2\omega_2(i)}{h^4} - \frac{\omega_1(i-h)}{h^2} \right] \\
&+ x_{i+2h} \left[\frac{\omega_2(i+h)}{h^4} \right]
\end{aligned}$$

8. References